

**Análisis multivariado para la identificación de componentes  
generadores de sabor y aroma en productos alimenticios**

Carolina Acosta Muñoz

Departamento de Ingeniería Eléctrica, Electrónica y Computación  
Facultad de Ingeniería y Arquitectura  
Universidad Nacional de Colombia Sede Manizales  
2010



**Análisis multivariado para la identificación de componentes  
generadores de sabor y aroma en productos alimenticios**

Carolina Acosta Muñoz

Trabajo de grado para optar al título de  
Magíster en Ingeniería — Automatización Industrial

Director

Prof. César Germán Castellanos Domínguez

Departamento de Ingeniería Eléctrica, Electrónica y Computación  
Facultad de Ingeniería y Arquitectura  
Universidad Nacional de Colombia Sede Manizales  
2010



**Multivariate data analysis with applications in food flavor  
compounds identification**

Carolina Acosta Muñoz

Thesis for the degree of  
Master in Engineering — Industrial Automation

Supervisor

Prof. César Germán Castellanos Domínguez

Department of Electrical, Electronic and Computer Engineering  
Faculty of Engineering and Architecture  
National University of Colombia at Manizales

2010



# Contenido

<b>Resumen</b>	ix
<b>Abstract</b>	xi
<b>1. Introducción</b>	1
<b>2. Objetivos</b>	5
<b>3. Marco teórico</b>	7
3.1. Datos Sensoriales	7
3.1.1. Análisis cuantitativo descriptivo	8
3.1.2. Diseño de experimentos	9
3.1.3. ANOVA	10
3.1.4. Análisis de componentes principales (PCA)	16
3.1.5. PCA-ANOVA	16
3.1.6. Corrección de escala Brockhoff	17
3.1.7. Corrección para diferentes usos de la escala entre los evaluadores	19
3.1.8. Análisis de Procrustes	20
3.1.9. Combinación de clasificadores	21
3.2. Cromatografía	22
3.2.1. Alineamiento dinámico en el tiempo (DTW)	23
3.2.2. Alineamiento de correlación optimizado (COW)	25
3.2.3. Algoritmo de ajuste de picos	26
3.3. Integración de datos cualitativos y cuantitativos	27
3.3.1. Regresión por mínimos cuadrados parciales (PLSR)	27
3.3.2. Análisis de correlación canónica (CCA)	29
3.4. Estado del arte	31
<b>4. Marco experimental</b>	35
4.1. Bases de datos	35
4.2. Protocolo de adquisición de datos sensoriales	35
4.2.1. Base de datos sensoriales e instrumentales de café	37
4.2.2. Base de datos sensoriales de papa	37
4.2.3. Base de datos sensoriales de queso	38
4.3. Análisis exploratorio de la uniformidad de los evaluadores	38
4.3.1. Interpretación de datos sensoriales usando análisis de varianza	39
4.3.2. Interpretación de datos sensoriales usando PCA-ANOVA	40
4.3.3. Evaluación de las correcciones del efecto de catador	43
4.4. Desempeño de los clasificadores	54
4.4.1. Verdaderos y falsos positivos y negativos	54
4.4.2. Medidas de desempeño	55

4.4.3. Validación cruzada	56
4.4.4. Comparación y evaluación mediante modelos discriminantes	56
4.4.5. Resultados para la base de datos de papas	62
4.4.6. Resultados de la base de datos de queso	65
4.5. Preprocesamiento de señales de cromatografía	66
4.5.1. Alineamiento de cromatografías	66
4.6. Integración de datos sensoriales y señales de cromatografía	69
4.6.1. Regresión por mínimos cuadrados parciales (PLS)	72
4.6.2. Correlación canónica (CCA)	75
<b>5. Discusión</b>	<b>79</b>
<b>6. Conclusión</b>	<b>81</b>
<b>A. ANOVA para cada atributo</b>	<b>83</b>
<b>Bibliografía</b>	<b>92</b>



## Resumen

Esta tesis de maestría es el resultado de mi proyecto de grado realizado en el Grupo de Control y Procesamiento Digital de Señales. El trabajo fue realizado durante el año 2009 y el primer semestre de 2010, bajo la supervisión del profesor César Germán Castellanos Domínguez del Departamento de Ingenierías Eléctrica, Electrónica y Computación.

Este trabajo está orientado a la construcción de una metodología de análisis multivariado, con el fin de encontrar patrones de asociación entre variables sensoriales e instrumentales medidas sobre muestras de café de origen, tostado a tres diferentes temperaturas y perteneciente a cuatro regiones de Colombia. Además, parte de la metodología se empleó para analizar datos sensoriales de otras aplicaciones de procesos industriales para productos alimenticios. Considerando lo anterior, es necesario el seguimiento de un protocolo de adquisición de datos sensoriales para procurar obtener una muestra lo más homogénea posible, ya que en este tipo de análisis las medidas pueden verse influenciadas por la subjetividad de cada miembro del panel sensorial.

En este trabajo se hizo especial énfasis a la etapa de preprocesamiento de los datos, debido a inconvenientes que resultaron durante el desarrollo del proyecto en la estandarización del procedimiento de evaluación sensorial. Este inconveniente fue abordado empleando transformaciones para corregir discrepancias entre miembros del panel de evaluaciones, así como una estrategia de combinación de clasificadores que aprovecha las experticias individuales. Los resultados muestran la ventaja de esta última estrategia en la identificación de clases.

Adicionalmente, en la etapa final del trabajo se derivan relaciones lineales entre el análisis sensorial e instrumental con el fin de descubrir dependencias de los atributos sensoriales sobre los compuestos químicos. La evidencia experimental muestra la influencia de factores como la temperatura de tostación del café y la dificultad en la caracterización de las regiones de origen de los productos alimenticios estudiados.



## Abstract

This Master's thesis is the result of my graduation project undertaken at the Control and Digital Signal Processing Group. The project was carried out during 2009 and the first semester of 2010, under the supervision of Prof. César Germán Castellanos-Domínguez from the Department of Electrical, Electronic and Computer Engineer.

This study is aimed at constructing a methodology of multivariate analysis, with the objective of finding association patterns between sensory variables and instrumental variables measured on samples of origin coffee, roasted at three different temperatures and belonging to four Colombian regions. Besides, parts of the methodology was used for analyzing sensory data from other industrial process applications for food products. According to that, it is mandatory to follow an acquisition protocol of sensory data in order to obtain a sample as homogeneous as possible since, in this type of analyses, measures may be influenced by subjectivity from each member of the sensory panel.

An special emphasis in the data preprocessing stage was done, due to shortcomings that appeared during the development of the project in the standarization of the sensory evaluation procedure. This shortcoming was faced by using transforms to correct discrepancies between panel members as well as by applying a classifier combination strategy that profits from individual expertises. Results show the advantage of the latter strategy for identifying classes.

In addition, in the final stage of the study, linear relationships between sensory analysis and instrumental analysis are extracted with the aim of finding dependencies of the sensory attributes and the chemical compounds. The experimental evidence shows the influence of factors such as the roasting temperature of the coffee and the difficulty for characterizing origin regions of the studied food products.



# 1. Introducción

Durante las últimas décadas un nuevo campo de investigación, conocido como quimiometría, ha ganado interés en el análisis de datos industriales. El propósito básico consiste en construir soluciones basadas en modelos matemáticos y estadísticos para analizar datos obtenidos de técnicas instrumentales como cromatografía de gases (GC), cromatografía líquida (LC) y espectrometría del infrarrojo cercano (NIR), entre otras. Adicionalmente, otras mediciones de gran interés en el ámbito industrial, especialmente en la industria alimenticia, son aquellas obtenidas por medio de análisis sensorial, el cual está estrechamente vinculado a la quimiometría desde que surgió la necesidad de fortalecer las evaluaciones sensoriales usando técnicas instrumentales.

En particular, la quimiometría comprende las tareas de recolección correcta de datos, optimización de parámetros experimentales, diseño de experimentos, calibración, procesamiento de señales y extracción de información relevante a partir de los datos instrumentales. Para implementar dichas tareas, se usan métodos estadísticos, técnicas de reconocimiento de patrones y de modelado matemático. Estas herramientas facilitan la ejecución de una subtarea importante: realizar análisis gráfico de los datos, mostrando la variación inherente a éstos y facilitando la interpretación. Algunas aplicaciones comprenden problemas como: regresión, para obtener la predicción de una variable dependiente a partir de ciertas variables independientes, medidas mediante procedimientos controlados donde quieran estudiarse los cambios generados bajo diferentes condiciones experimentales; correlaciones simples y múltiples, que indican dependencia entre variables que pueden o no ser de la misma naturaleza; clasificación, que permite obtener un modelo que sirva para caracterizar grupos de interés, determinando sus diferencias con el fin de tomar una decisión sobre la pertenencia de una muestra a cierto grupo. La importancia del uso de modelos estadísticos sobre datos químicos está en poder estudiar patrones inesperados debidos al efecto conjunto de todas las variables, en contraste con las relaciones químicas y físicas tradicionales que normalmente consideran solo una o muy pocas variables al mismo tiempo [Wold, 1995]. Además, en el enfoque quimiométrico se tienen en cuenta y se cuantifican la presencia de ruido, la colinealidad y los errores en los modelos.

El área de quimiometría aún tiene muchos retos, especialmente por la naturaleza de los diferentes tipos de datos — espectrales, medidas físicas, evaluaciones sensoriales— que deben estudiarse. Además, los datos instrumentales generalmente implican alta dimensión y multicolinealidad, lo cual, a su vez, se traduce en un mal condicionamiento del problema al tener muchas más variables que observaciones. Esta situación es difícil de superar debido al alto costo, en tiempo y dinero, de generar un número de observaciones mayor que el número de variables medidas, por ejemplo, un espectro de cromatografía siempre tiene cientos o miles

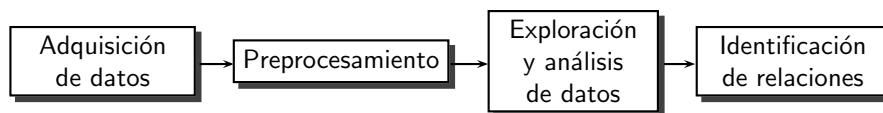
de puntos que se consideran individualmente como variables; más aún, la obtención de la cromatografía de una muestra puede tardar más 30 minutos.

En particular, en la investigación y el desarrollo de alimentos, es frecuente emplear análisis sensorial como medio para determinar la calidad de producción de un alimento; así como para generar nuevos productos, y seleccionar las materias prima. Sin embargo, esta práctica está basada en el uso de los sentidos humanos como instrumentos de medición, lo cual, inevitablemente introduce subjetividades en las evaluaciones. El estudio de evaluaciones sensoriales ha sido objeto de investigación durante décadas [Moskowitz, 1996]. No obstante, la selección de evaluadores, el entrenamiento del panel sensorial, la determinación del desempeño grupal e individual, la reproducibilidad y el seguimiento y mejoramiento del desempeño del panel son aún problemas abiertos de investigación [Kermit and Lengard, 2005].

En esta tesis se aborda el tema de encontrar asociaciones entre datos sensoriales e instrumentales obtenidos por medio de cromatografía de gases, particularmente en el marco del proyecto cuyo objetivo es caracterizar ciertas regiones del país con respecto a sus aromas. El estudio de las propiedades sensoriales basado en medidas instrumentales permite identificar cuáles de las variables instrumentales influyen de mayor forma en la generación de aromas, sabores, texturas, etc., de este modo, encontrando relaciones existentes se podría suministrar soporte automático a los análisis sensoriales y a largo plazo lograr determinar propiedades sensoriales únicamente mediante el estudio de sus medidas instrumentales al haber determinado con precisión las relaciones. Con base en lo anterior se hace necesario establecer un método que permita analizar de forma simultánea las variables disponibles tanto sensoriales como instrumentales en el proceso de entrenamiento de sistemas.

El tema anterior está motivado por el interés de la industria colombiana de café liofilizado de analizar y determinar los componentes generadores de aroma de los cafés de origen, de tal forma que se logre caracterizar el aroma del café de cada región. Esta caracterización permitiría generar nuevos productos así como verificar la autenticidad de la materia prima.

FIGURA 1.1. Diagrama de bloques del sistema de análisis de datos



El diagrama de la figura 1.1 muestra las etapas generales de un sistema de identificación de relaciones entre datos sensoriales e instrumentales. La primera etapa incluye un procedimiento estandarizado de adquisición de datos, tanto el protocolo de evaluación sensorial como de la calibración de la adquisición instrumental, donde la cromatografía de gases proporciona los compuestos químicos presentes en el producto. La segunda etapa, con respecto

al análisis sensorial, corresponde al estudio de concordancia entre los evaluadores, la corrección de discrepancias individuales y el aprovechamiento de las habilidades individuales en la toma conjunta de decisiones; respecto a los datos cromatográficos, esta etapa abarca el problema de desalineamiento entre las señales medidas. En la etapa de exploración y análisis de datos, se emplearon gráficas de dispersión y experimentos de clasificación para comprobar la existencia o ausencia de agrupaciones naturales. También en esta etapa se emplearon técnicas de reducción de dimensión y selección de variables. En la etapa de identificación de relaciones se realiza un análisis integrado por los dos tipos de variables, sensoriales e instrumentales, con el fin de determinar los compuestos químicos que generan los atributos sensoriales.

Este trabajo está organizado como sigue. En el capítulo 2 se presentan los objetivos de la propuesta de tesis. En el capítulo 3 se presentan las técnicas de análisis estadístico y de procesamiento de señales empleadas a lo largo del trabajo. El capítulo 4 se presentan las bases de datos analizadas, la metodología de los experimentos, los resultados y su respectiva discusión. En el capítulo 5 se resaltan las observaciones más importantes de los experimentos. Finalmente, en el capítulo 6 se resumen algunas observaciones concluyentes, se propone trabajo a futuro y se discuten los aportes realizados en el desarrollo de este trabajo.

La recolección o adquisición de los datos sensoriales requieren un especial cuidado y seguimiento de un protocolo estandarizado (sección 4.2). Sin embargo, a menudo las recomendaciones de las normas técnicas no se obedecen estrictamente. De hecho, en el caso de los conjuntos de datos obtenidos en el marco de la cooperación con la industria de café infortunadamente se desatendieron estas recomendaciones debido a la existencia de un protocolo de catación propio de la empresa. Algunas de las discrepancias entre lo recomendado por la norma y la practica de catación realizada, incluyen: la selección de atributos, escala de medición, entrenamiento adecuado, etc. Estas inconsistencias no pudieron ser corregidas con posterioridad debido a la presuposición de que el panel de catadores profesionales seguía el protocolo estandarizado.





## 2. Objetivos

### Objetivo general

Construir una metodología de análisis multivariado, que permita encontrar los respectivos patrones de asociación entre las variables sensoriales e instrumentales, que pueda ser aplicada a procesos industriales de productos alimenticios.

### Objetivos específicos

- Definición de un protocolo a seguir para determinar las unidades experimentales y realizar el análisis sensorial e instrumental.
- Desarrollo de una metodología de preprocesamiento de las medidas de análisis sensorial e instrumental que garantice la validez de los resultados.
- Construcción de una metodología de identificación y extracción de las variables que sean relevantes en la asociación de variables.
- Identificación, implementación, evaluación y comparación de las diferentes técnicas para el análisis de las relaciones entre las variables sensoriales que describan adecuadamente los atributos de fragancia del producto y las variables instrumentales que describan la composición del producto.



## 3. Marco teórico

### 3.1. Datos Sensoriales

La caracterización de productos basada en evaluaciones sensoriales juega un papel importante en el control de calidad. Uno de los problemas del análisis de perfiles sensoriales es tener que manejar las diferencias sistemáticas individuales en las evaluaciones de los panelistas. Es inevitable que cada evaluador tenga, en cierto grado, diferentes percepciones de las muestras así como diferentes apreciaciones de los atributos o de las escalas usadas para cuantificar las mediciones. La forma de evaluación de los panelistas no es lineal, ya que no se puede garantizar que a un cambio de intensidad de sabor de cierta magnitud (en cualquier parte de la escala de sabores / olores) corresponda siempre a un cambio en la evaluación numérica de la misma proporción. Es decir, para los sabores fuertes, los catadores pueden tender a evaluar los pequeños cambios con grandes cambios en la escala de evaluación; mientras que los grandes cambios para los sabores débiles pueden llegar a ser evaluados con pequeños cambios en la escala de evaluación. En otras palabras, la respuesta de los evaluadores puede saturarse en alguno de los dos extremos de la escala, dependiendo de su sensibilidad. Por lo tanto, cualquier modelo que pretenda describir perfiles sensoriales necesita considerar las percepciones individuales; implícita o explícitamente. Para un único atributo sensorial, las diferencias fundamentales entre catadores pueden dividirse en diferencias en (i) nivel de la escala, (ii) sensibilidad, (iii) desacuerdo y (iv) variabilidad [Brockhoff, 2003]. Es importante investigar las similitudes y diferencias entre los objetos, catadores y atributos al mismo tiempo.

El tipo de preguntas que surgen al analizar este tipo de datos son: ¿Los evaluadores usan los atributos o las escalas de medida de forma diferente?, ¿algunos evaluadores son más sensibles que otro a algunos atributos?, ¿algunos evaluadores son mejores identificando diferencias entre cierto grupo de productos?, ¿todos los evaluadores distinguen igualmente bien los productos?, ¿todos los evaluadores usan los mismos atributos para distinguir los productos? Todas estas preguntas son de interés para el líder del panel, quien es responsable de la calidad del mismo y podría decidir si se conserva o se prescinde de alguno de los evaluadores. Igualmente, los mismos interrogantes son de interés para el analista de los datos que debe tomar decisiones sobre cuál técnica es la más apropiada para caracterizar los diferentes productos. Algunos métodos de corrección de escala encontrados en la literatura como el modelo del evaluador [Brockhoff and Skovgaard, 1994], análisis de Procrustes [Dijksterhuis, 1996] y el método de Ten Berge [Romano et al., 2008], buscan corregir las evaluaciones sensoriales eliminando las diferencias individuales y homogeneizando los datos para conseguir una caracterización de los productos independiente de la capacidad individual de los evaluadores. Por otro lado, en [Granitto et al., 2008] se propone aprovechar las diferencias individuales de los evaluadores obteniendo un modelo de caracterización de los productos

para cada evaluador por separado, y mediante el empleo de un ensamble de clasificadores, mejorar la identificación de los productos.

En esta etapa de la tesis se propone comparar estas dos metodologías en la identificación de productos alimenticios usando datos sensoriales. La eficacia de los tratamientos de corrección de escala se mide mediante el empleo de pruebas de hipótesis usando ANOVA de dos y tres vías. Se pretende estudiar el efecto de los evaluadores, en cuanto a su percepción individual del nivel y la escala, así como los efectos del producto y de la interacción entre catador y producto, en cuanto a la percepción individual que los evaluadores tienen de cada producto. Para comparar la metodología de ensamble de clasificadores con las de corrección de escala, se emplea la precisión de clasificación de los productos usando clasificadores sencillos, como el clasificador del vecino más cercano y el clasificador discriminante lineal.

### **3.1.1. Análisis cuantitativo descriptivo**

El análisis cuantitativo descriptivo (QDA por sus siglas en inglés) es una de las técnicas principales de análisis descriptivo para evaluaciones sensoriales. La intención principal de este método es manejar problemas de muestra pequeña, frecuentemente obtenidos en aplicaciones de perfiles de sabor y aroma. Este tipo de análisis sensorial es una técnica reproducible para identificar, cuantificar y describir las características de un producto y determinar su calidad sensorial; éste es un método que permite evaluar tanto individual como integralmente las características sensoriales de cierto producto y las relaciones entre ellas, obteniéndose la mayor información posible de una muestra [Meilgaard et al., 2007]. En este análisis se evalúan características de la apariencia, aroma, sabor y textura de un producto tanto de sus atributos como defectos y se emite una calificación integral que resume y agrupa el concepto del juez entrenado sobre la calidad de la muestra.

Los catadores son seleccionados de un gran conjunto de posibles candidatos de acuerdo a su habilidad para discriminar diferencias en los perfiles sensoriales entre las muestras de un tipo específico de producto. Posteriormente, los catadores deben ser sometidos a un entrenamiento constante y exhaustivo, con el fin de garantizar el análisis con catadores entrenados y/o expertos que permita obtener resultados confiables para una adecuada interpretación. El entrenamiento de un panel de QDA requiere el uso de referencias de los productos y sus ingredientes para estimular la generación de la terminología consistente. El líder del panel actúa como un facilitador, más que como un instructor, y debe abstenerse de influenciar al grupo. Las calificaciones se realizan frecuentemente usando una línea de escala de 15 cm. Un panel debidamente entrenado tiene la capacidad de cuantificar diferencias entre las muestras, evaluando diferentes características o propiedades sensoriales en una escala de intervalos estructurada.

En la metodología QDA, se proponen evaluaciones para múltiples productos con el fin de sacar el mayor provecho de las capacidades de los evaluadores para hacer juzgamiento con un alto grado de precisión. Los evaluadores son buenos para juzgar diferencias sensoriales relativas, dependientes de su experiencia, pero no puede esperarse obtener diferencias absolutas. Se debe tener en cuenta que el entrenamiento y la selección adecuados del panel

de catación es determinante para lograr un análisis confiable de los datos y del producto evaluado. La falta de instrucción formal puede conducir al desarrollo de términos erróneos. Sin embargo, es reconocido y aceptado en QDA que los panelistas pueden usar diferentes partes de la escala para expresar sus percepciones de la misma muestra.

En QDA, los panelistas evalúan los productos uno a la vez en cabinas separadas para reducir las distracciones y la interacción entre panelistas. No está permitido que los panelistas discutan sobre los resultados, la terminología o la muestra después de cada sesión y el líder del panel debe guardar discreción sobre cualquier diferencia entre las muestras.

Los resultados de QDA son adecuados para los análisis estadísticos y el reporte general debe contener representaciones gráficas de los datos. El desempeño del panel puede ser examinado usando un análisis de interacción entre el producto y el panelista; las diferencias entre los productos puede diagnosticarse usando procedimientos tales como Análisis de varianza — univariado y multivariado—, análisis de componentes principales, análisis factorial y análisis de conglomerados, entre otros.

### 3.1.2. Diseño de experimentos

Un diseño estadístico de experimentos es el procedimiento de planear los experimentos tal que los datos recolectados sean apropiados para ser analizados estadísticamente y se conduzca a conclusiones objetivas. En esta sección se presenta el diseño de experimentos factorial, en el cual las corridas del experimento se ejecutan con todas las combinaciones posibles de los niveles de los factores y de esta forma poder estudiar la interacción entre estos últimos.

La interacción entre los factores se debe tener en cuenta ya que permite explicar la respuesta del fenómeno con más elementos de juicio. Gracias a las interacciones, en el caso de dos factores, es posible determinar si un factor modifica el efecto de otro factor sobre la variable respuesta y en el caso de tres factores, determinar si un factor modifica la interacción de los otros dos factores.

Las observaciones, de un experimento factorial de dos factores, pueden describirse con el modelo estadístico lineal

$$Y_{ijr} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijr} \quad \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, b \\ r = 1, 2, \dots, n \end{cases} \quad (1)$$

donde  $\mu$  es el efecto promedio global,  $\tau_i$  es el efecto del  $i$ -ésimo nivel del factor A,  $\beta_j$  es el efecto del  $j$ -ésimo nivel del factor B,  $(\tau\beta)_{ij}$  es el efecto de la interacción entre A y B y  $\epsilon_{ijr}$  es un componente de error aleatorio que tiene una distribución normal con media cero y varianza  $\sigma^2$ . Además, cada factor A y B tiene a y b niveles, respectivamente.

Otro posible diseño de experimentos se obtiene al tener en cuenta tres factores y se puede describir el modelo subyacente como sigue

$$Y_{ijk r} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{ik} + (\tau\beta\gamma)_{ijk} + \epsilon_{ijk r} \quad \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, b \\ k = 1, 2, \dots, c \\ r = 1, 2, \dots, n \end{cases} \quad (2)$$

donde los efectos  $\mu$ ,  $\tau_i$ ,  $\beta_j$  y  $\gamma_k$  corresponden a los efectos principales y  $(\tau\beta)_{ij}$ ,  $(\tau\gamma)_{ik}$ ,  $(\beta\gamma)_{ik}$  y  $(\tau\beta\gamma)_{ijk}$  corresponden a la interacción entre los diferentes efectos.

La prueba de hipótesis consiste en probar que no existen efectos del factor A, del factor B y del factor C sobre la respuesta y además de que no existe interacción entre los factores A, B y C, por parejas o en conjunto. Esta hipótesis es probada haciendo uso del análisis de varianza (ANOVA) de dos factores y tres factores. Antes de iniciar el ANOVA para estos diseños de experimentos es importante determinar si los factores son de tipo fijo o aleatorio. Un factor es fijo cuando los niveles de cierto factor, son seleccionados de manera específica por el experimentador, de tal forma que las inferencias están limitadas únicamente a estos niveles. En lo que se refiere a factores aleatorios, éstos corresponden a situaciones donde existen una gran cantidad de niveles posibles para el factor de interés, y el experimentador debe escoger al azar de toda la población un subconjunto de estos niveles, tal que, a partir de este subconjunto, se puedan obtener conclusiones para todos los posibles niveles del factor [Montgomery and Runger, 2002].

### 3.1.3. ANOVA

El propósito de ANOVA es probar significancia estadística de las diferencias entre las medias de los grupos o niveles de un factor. Esto se logra analizando la varianza, es decir, partiendo la varianza total en el componente que se debe a un error aleatorio de la medida y el componente que se debe a las diferencias entre las medias de los niveles del factor. Finalmente, estas varianzas son comparadas para probar la hipótesis nula de que no hay diferencia entre las medias, y en caso de que resulten significativamente diferentes, se rechaza dicha hipótesis.

TABLA 3.1. Arreglo de datos para un diseño factorial con dos factores

		Factor B				Totales	Promedios
		1	2	...	b		
Factor A	1	$y_{111}, y_{112},$ $\dots, y_{11n},$	$y_{121}, y_{122},$ $\dots, y_{12n},$		$y_{1b1}, y_{1b2},$ $\dots, y_{1bn},$	$y_{1..}$	$\bar{y}_{1..}$
	2	$y_{211}, y_{212},$ $\dots, y_{21n},$	$y_{221}, y_{222},$ $\dots, y_{22n},$		$y_{2b1}, y_{2b2},$ $\dots, y_{2bn},$	$y_{2..}$	$\bar{y}_{2..}$
	⋮						
	a	$y_{a11}, y_{a12},$ $\dots, y_{a1n},$	$y_{a21}, y_{a22},$ $\dots, y_{a2n},$		$y_{ab1}, y_{ab2},$ $\dots, y_{abn},$	$y_{a..}$	$\bar{y}_{a..}$
Totales		$y_{.1.}$	$y_{.2.}$		$y_{.b.}$	$y_{...}$	
Promedios		$\bar{y}_{.1.}$	$\bar{y}_{.2.}$		$\bar{y}_{.b.}$		$\bar{y}_{...}$

En la tabla 3.1 se muestra la recopilación de los datos medidos para un diseño de experimentos de dos factores, que así mismo puede extenderse al caso de tres factores. De acuerdo a la tabla se definen  $y_{i..}$  como la suma total de las observaciones tomadas en el  $i$ -ésimo nivel del factor  $A$ ;  $y_{.j.}$  la suma total de observaciones tomadas en el  $j$ -ésimo nivel del factor  $B$ ;  $y_{ij.}$  la suma total de las observaciones en la  $ij$ -ésima celda de la tabla 3.1, y  $y$  el gran total de todas las observaciones. Defínanse  $\bar{y}_{i..}$ ,  $\bar{y}_{.j.}$ ,  $\bar{y}_{ij.}$ , y  $\bar{y}_{..}$  como los promedios por renglón, columna, celda y total, respectivamente.

Para probar la diferencia estadística entre los niveles de un factor se tienen en cuenta tres casos, se prueba el efecto de los factores individualmente, se prueba el efecto de la interacción entre los pares de factores posibles y se prueba el efecto de la interacción entre los tres factores presentes en el estudio.

Para el primer caso, cada factor individual debe ser evaluado para probar su efecto sobre la medición  $Y_{ijk\tau}$  (ver ecuación (2)). En particular, para el factor  $\tau$  la hipótesis nula, denotada como  $H_0$ , consiste en que todos los  $a$  niveles del factor son iguales a 0. O sea que no hay diferencias significativas entre dichos niveles. Mientras que si al menos para un nivel  $i$  no se cumple esta igualdad, se rechaza la hipótesis nula y se acepta la hipótesis alternativa  $H_1$ . Una descripción similar puede hacerse para los demás factores tenidos en cuenta. A continuación se presenta la notación habitual de las hipótesis a probar para cada uno los factores principales  $\tau$ ,  $\beta$ ,  $\gamma$ :

1.  $H_0 : \tau_1 = \tau_2 = \dots = \tau_a = 0$   
 $H_1 : \text{al menos para un nivel } i \tau_i \neq 0$
2.  $H_0 : \beta_1 = \beta_2 = \dots = \beta_b = 0$   
 $H_1 : \text{al menos para un nivel } j \beta_j \neq 0$
3.  $H_0 : \gamma_1 = \gamma_2 = \dots = \gamma_c = 0$   
 $H_1 : \text{al menos para un nivel } k \gamma_k \neq 0$

Para el segundo caso, se prueba el efecto de la interacción entre cada par de factores sobre la medición  $Y_{ijk\tau}$ . Los posibles pares entre los factores  $\tau$ ,  $\beta$  y  $\gamma$  corresponden a  $\tau\beta$  con  $ij$  niveles,  $\tau\gamma$  con  $ik$  niveles y  $\beta\gamma$  con  $jk$  niveles. En particular, para el factor de interacción  $\tau\beta$ , la hipótesis nula  $H_0$  expresa que no hay diferencias entre los  $ab$  niveles de interacción entre los factores  $\tau$  y  $\beta$ . Si alguno de los niveles no cumple esta hipótesis se rechaza la hipótesis nula y se acepta la hipótesis alternativa  $H_1$ . Las hipótesis de los demás factores de interacción se describen de forma similar. A continuación se presentan las pruebas hipótesis de interacción que se tienen en cuenta en un diseño de experimentos con tres factores.

4.  $H_0 : (\tau\beta)_{11} = (\tau\beta)_{12} = \dots = (\tau\beta)_{ab} = 0$   
 $H_1 : \text{al menos para un par de niveles } ij \ (\tau\beta)_{ij} \neq 0$
5.  $H_0 : (\tau\gamma)_{11} = (\tau\gamma)_{12} = \dots = (\tau\gamma)_{ac} = 0$   
 $H_1 : \text{al menos para un par de niveles } ik \ (\tau\gamma)_{ik} \neq 0$

6.  $H_0 : (\beta\gamma)_{11} = (\beta\gamma)_{12} = \dots = (\beta\gamma)_{bc} = 0$   
 $H_1 : \text{al menos para un par de niveles } jk \text{ } (\beta\gamma)_{jk} \neq 0$

El tercer caso consiste en probar la interacción entre los tres factores. La hipótesis nula  $H_0$  consiste en probar que no hay diferencias entre los  $abc$  niveles de interacción. Si no se cumple dicha hipótesis se acepta la hipótesis alternativa  $H_1$  de que el efecto para al menos una combinación de niveles  $ijk$  no es igual a los demás efectos.

7.  $H_0 : (\tau\beta\gamma)_{111} = (\tau\beta\gamma)_{112} = \dots = (\tau\beta\gamma)_{abc} = 0$   
 $H_1 : \text{al menos para una combinación de niveles } ijk \text{ } (\tau\beta\gamma)_{ijk} \neq 0$

La variabilidad total se mide con la suma total de cuadrados de las observaciones así

$$SS_T = \sum_{i=1}^a \sum_{j=1}^b \sum_{r=1}^n (y_{ijr} - \bar{y}_{...})^2 \quad (3)$$

mientras que la descomposición de la variabilidad está definida de la siguiente manera para un análisis de varianza de dos factores

$$\begin{aligned} \sum_{i=1}^a \sum_{j=1}^b \sum_{r=1}^n (y_{ijr} - \bar{y}_{...})^2 &= bn \sum_{i=1}^a (\bar{y}_{i..} - \bar{y}_{...})^2 \\ &+ an \sum_{j=1}^b (\bar{y}_{.j.} - \bar{y}_{...})^2 \\ &+ n \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 \\ &+ \sum_{i=1}^a \sum_{j=1}^b \sum_{r=1}^n (y_{ijr} - \bar{y}_{ij.})^2 \end{aligned} \quad (4)$$

Esta suma de cuadrados puede ser fácilmente extendida al caso de tres factores. Es así como la suma total de cuadrados queda dividida en una suma de cuadrados debida al factor  $A$  ( $SS_A$ ), una suma de cuadrados debida al factor  $B$  ( $SS_B$ ), una suma de cuadrados debida a la interacción entre  $A$  y  $B$  ( $SS_{AB}$ ) y una suma de cuadrados debida al error ( $SS_E$ ).

La identidad de la suma de cuadrados de la ecuación 4 puede escribirse de manera simbólica como

$$SS_T = SS_A + SS_B + SS_{AB} + SS_E \quad (5)$$

Existen en total  $abn - 1$  grados de libertad. Los efectos principales  $A$  y  $B$  tienen  $a - 1$  y  $b - 1$  grados de libertad, mientras que el efecto de interacción  $AB$  tiene  $(a - 1)(b - 1)$  grados de libertad. Dentro de cada una de las  $ab$  celdas de la tabla 3.1, existen  $n - 1$  grados de libertad entre las  $n$  réplicas, y las observaciones en la misma celda sólo pueden diferir debido al error aleatorio. Por consiguiente, existen  $ab(n - 1)$  grados de libertad para



el error. En consecuencia, los grados de libertad están divididos de acuerdo con

$$abn - 1 = (a - 1) + (b - 1) + (a - 1)(b - 1) + ab(n - 1) \quad (6)$$

Si se divide cada una de las sumas de cuadrados del lado derecho de la ecuación 5 entre el número de grados de libertad correspondiente, se obtienen las medias de cuadrados de  $A$ ,  $B$ , la interacción y el error:

$$MS_A = \frac{SS_A}{a - 1} \quad MS_B = \frac{SS_B}{b - 1}$$

$$MS_{AB} = \frac{SS_{AB}}{(a - 1)(b - 1)} \quad MS_E = \frac{SS_E}{ab(n - 1)}$$

La variabilidad ( $SS_E$ ) dentro del grupo se conoce como la varianza del error. Este término denota el hecho de que esta variabilidad no podemos explicarla fácilmente en el diseño actual. Sin embargo, la variabilidad  $SS_A$ ,  $SS_B$ ,  $SS_{AB}$  de los efectos sí se pueden explicar. En efecto, éstas se deben a las diferencias en las medias entre los grupos. En otras palabras, la pertenencia a una clase explica la variabilidad porque se sabe de antemano que ésta se debe a las diferencias entre las medias.

Para probar que los efectos de los niveles para cierto factor son todos iguales a cero ( $H_0 : \tau_i = 0$ ), se emplea el cociente

$$F_0 = \frac{MS_A}{MS_E} \quad (7)$$

el cual tiene una distribución F con  $a - 1$  y  $ab(n - 1)$  grados de libertad si  $H_0 : \tau_i = 0$  es verdadera. Esta hipótesis nula se rechaza con un nivel de significancia  $\alpha$  si  $f_0 > f_{\alpha, a-1, ab(n-1)}$ . De manera similar se define el cociente, la prueba de hipótesis y los grados de libertad para los demás efectos e interacciones.

Las pruebas de hipótesis se basan en la comparación de las varianzas debido a la variabilidad dentro de los grupos ( $MS_A$ ) con respecto a la variabilidad entre grupos ( $MS_E$ ). Bajo la hipótesis nula, de que no hay diferencias en la media entre grupos para cierta población, aún se esperaría alguna fluctuación aleatoria en las medias para los dos grupos cuando se toma una muestra pequeña. Por lo tanto, bajo la hipótesis nula, la varianza estimada con base en la variabilidad dentro de los grupos debe ser la misma que la varianza debido a la variabilidad entre grupos. Se puede comparar esos dos estimadores de varianza usando la prueba F, que prueba en qué casos esa relación de las dos estimaciones de varianza es significativamente más grande que 1. En el ejemplo, la prueba es significativamente mayor, y se concluye que las medias para los dos grupos son significativamente diferentes.

Normalmente los cálculos se muestran en una tabla de análisis de varianza tal como se muestra en las tablas 3.2 y 3.3

TABLA 3.2. Tabla de análisis de varianza para el modelo de efectos fijos de dos factores

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados	Medias de cuadrados esperadas	F
A	$SS_A$	$a - 1$	$MS_A$	$\sigma^2 + \frac{bcn\Sigma\tau_i^2}{a - 1}$	$\frac{MS_A}{MS_E}$
B	$SS_B$	$b - 1$	$MS_B$	$\sigma^2 + \frac{acn\Sigma\beta_j^2}{b - 1}$	$\frac{MS_B}{MS_E}$
AB	$SS_{AB}$	$(a - 1)(b - 1)$	$MS_{AB}$	$\sigma^2 + \frac{cn\Sigma\Sigma(\tau\beta)_{ij}^2}{(a - 1)(b - 1)}$	$\frac{MS_{AB}}{MS_E}$
Error	$SS_E$	$ab(n - 1)$	$MS_E$	$\sigma^2$	
Total	$SS_T$	$abn - 1$			

TABLA 3.3. Tabla de análisis de varianza para el modelo de efectos fijos de tres factores

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados	Medias de cuadrados esperadas	F
A	$SS_A$	$a - 1$	$MS_A$	$\sigma^2 + \frac{bcn\Sigma\tau_i^2}{a - 1}$	$\frac{MS_A}{MS_E}$
B	$SS_B$	$b - 1$	$MS_B$	$\sigma^2 + \frac{acn\Sigma\beta_j^2}{b - 1}$	$\frac{MS_B}{MS_E}$
C	$SS_C$	$c - 1$	$MS_C$	$\sigma^2 + \frac{abn\Sigma\gamma_k^2}{c - 1}$	$\frac{MS_C}{MS_E}$
AB	$SS_{AB}$	$(a - 1)(b - 1)$	$MS_{AB}$	$\sigma^2 + \frac{cn\Sigma\Sigma(\tau\beta)_{ij}^2}{(a - 1)(b - 1)}$	$\frac{MS_{AB}}{MS_E}$
AC	$SS_{AC}$	$(a - 1)(c - 1)$	$MS_{AC}$	$\sigma^2 + \frac{bn\Sigma\Sigma(\tau\gamma)_{ik}^2}{(a - 1)(c - 1)}$	$\frac{MS_{AC}}{MS_E}$
BC	$SS_{BC}$	$(b - 1)(c - 1)$	$MS_{BC}$	$\sigma^2 + \frac{an\Sigma\Sigma(\beta\gamma)_{jk}^2}{(b - 1)(c - 1)}$	$\frac{MS_{BC}}{MS_E}$
ABC	$SS_{ABC}$	$(a - 1)(b - 1)(c - 1)$	$MS_{ABC}$	$\sigma^2 + \frac{n\Sigma\Sigma\Sigma(\tau\beta\gamma)_{ijk}^2}{(a - 1)(b - 1)(c - 1)}$	$\frac{MS_{ABC}}{MS_E}$
Error	$SS_E$	$abc(n - 1)$	$MS_E$	$\sigma^2$	
Total	$SS_T$	$abcn - 1$			

## MANOVA

El modelo ANOVA multivariado (MANOVA) de un factor se expresa como:

$$\begin{aligned} \mathbf{y}_{ir} &= \boldsymbol{\mu} + \boldsymbol{\alpha}_i + \boldsymbol{\varepsilon}_{ir} \\ &= \boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_{ir}, \quad i = 1, 2, \dots, a \quad r = 1, 2, \dots, n \end{aligned} \quad (8)$$

en términos de las  $p$  variables en  $\mathbf{y}_{ir}$ , (8) se vuelve

$$\begin{pmatrix} y_{ir1} \\ y_{ir2} \\ \vdots \\ y_{irp} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} + \begin{pmatrix} \alpha_{i1} \\ \alpha_{i2} \\ \vdots \\ \alpha_{ip} \end{pmatrix} + \begin{pmatrix} \varepsilon_{ir1} \\ \varepsilon_{ir2} \\ \vdots \\ \varepsilon_{irp} \end{pmatrix} = \begin{pmatrix} \mu_{i1} \\ \mu_{i2} \\ \vdots \\ \mu_{ip} \end{pmatrix} + \begin{pmatrix} \varepsilon_{ir1} \\ \varepsilon_{ir2} \\ \vdots \\ \varepsilon_{irp} \end{pmatrix} \quad (9)$$

de tal forma que el modelo para la  $l$ -ésima variable (para  $l = 1, 2, \dots, p$ ) en cada vector  $\mathbf{y}_{ij}$  es

$$y_{irl} = \mu_l + \alpha_{il} + \varepsilon_{irl} = \mu_{il} + \varepsilon_{irl} \quad (10)$$

Se desea comparar los vectores del factor  $\boldsymbol{\mu}$  sobre los  $a$  niveles para encontrar diferencias significativas. Por lo tanto, la prueba de hipótesis se expresa

$$\begin{aligned} H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \dots = \boldsymbol{\mu}_a = \mathbf{0} \\ H_1 : \text{al menos para un nivel } i \quad \boldsymbol{\mu}_i \neq \mathbf{0} \end{aligned} \quad (11)$$

La igualdad de los vectores del factor  $\boldsymbol{\mu}$  implica que los  $a$  niveles son iguales para cada variable; es decir,  $\mu_{1l} = \mu_{2l} = \dots = \mu_{al}$  para  $l = 1, 2, \dots, p$ . Si dos pares de niveles para cierta variable son diferentes, por ejemplo,  $\mu_{23} \neq \mu_{43}$ , entonces  $H_0$  es falsa y es rechazada. En términos matriciales la hipótesis nula se expresa:

$$H_0 : \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \vdots \\ \mu_{1p} \end{pmatrix} = \begin{pmatrix} \mu_{21} \\ \mu_{22} \\ \vdots \\ \mu_{2p} \end{pmatrix} = \dots = \begin{pmatrix} \mu_{a1} \\ \mu_{a2} \\ \vdots \\ \mu_{ap} \end{pmatrix} \quad (12)$$

Siguiendo estos criterios básicos, los modelos ANOVA para tres factores se pueden expresar de forma multivariada, para obtener un diseño de experimentos multivariado con el número de factores y niveles requeridos.

### La prueba de diferencias de Tukey HSD

Una prueba de significancia F dice solo que la diferencia entre las medias de varias muestras o grupos es significativamente mayor que cero. Pero no dice en qué circunstancias la media de un grupo particular es significativamente diferente de otro grupo. La prueba de diferencias de Tukey HSD (HSD de su nombre en inglés *honestly significant difference*) permite determinar la significancia de las diferencias por parejas entre los grupos.

El procedimiento de comparación múltiple de Tukey consiste en la estimación, entre todas las posibles parejas de medias de niveles para cada factor, de las diferencias  $D_i = \mu_i - \mu'_i$ , o la estimación de todas las pruebas de la forma:

$$\begin{aligned} H_0 : \mu_i - \mu_{i'} &= 0 \\ H_a : \mu_i - \mu_{i'} &\neq 0 \end{aligned}$$

Cuando todos los grupos son de igual tamaño, el coeficiente de confianza del método de Tukey es de exactamente  $1 - \alpha$  y los niveles de significancia es exactamente  $\alpha$ . Cuando los tamaños de los grupos no son iguales, el coeficiente de confianza es mayor que  $1 - \alpha$  y el nivel de significancia es menor que  $\alpha$ . En otras palabras, el procedimiento de Tukey es conservador cuando el tamaño de los grupos no son iguales [Kutner et al., 2005]. La prueba de Tukey utiliza una *distribución de rango estudentizada*, que puede abreviarse como  $q$ . Para cualquier par de medias entre  $k$  grupos, se determinan la mayor y menor como  $M_L$  y  $M_S$ , respectivamente. El estadístico de rango estudentizado puede ser calculado para cualquier par de medias como

$$q = \frac{M_L - M_S}{\sqrt{MS_E/N_i}} \quad (13)$$

donde  $MS_E$  es la media de cuadrados del error o varianza intra grupo obtenida en el análisis original y  $N_i$  es el número de observaciones del grupo  $i$ .

### 3.1.4. Análisis de componentes principales (PCA)

Sea  $\mathbf{Y}$  la matriz original de datos con dimensiones  $m \times p$ . Las filas  $y_i$  corresponden a las observaciones y las columnas  $\xi_j$  a las variables. Las variables son estandarizadas con media cero y varianza uno (i.e.,  $\mathbf{Y} = (\xi_l - \bar{\xi}_l)/s_{\xi_l}$ ), tomando a  $s_{\xi_l}$  como la desviación estándar de  $\xi_l$ ). La matriz estandarizada  $\mathbf{Y}$  queda representada como

$$\mathbf{Y} = \{y_{il} : i = 1, \dots, m \quad l = 1, \dots, p\} \quad (14)$$

con vector de medias  $\bar{\xi}_l = E\{\xi_l\}$  y matriz de covarianza  $\Sigma_{\mathbf{Y}} = E\{(\xi - \mu_{\xi})^T(\xi - \mu_{\xi})\} > 0$ , de forma que algunos de sus elementos o componentes están correlacionados. La ecuación matricial de PCA puede expresarse como

$$\mathbf{T} = \mathbf{Y}\mathbf{P} \quad (15)$$

donde  $\mathbf{T}$  es la matriz de componentes principales,  $\mathbf{Y}$  es la matriz de datos estandarizada y  $\mathbf{P}$  es la matriz de transformación formada por los vectores propios correspondientes a cada valor propio  $\lambda_j$  de la matriz de covarianza  $\Sigma_{\mathbf{Y}}$ . Estos resultados son representados en diagramas de dos y tres dimensiones donde los ejes coordenados son formados por las componentes principales y la ubicación de las variables sobre el diagrama está establecida por los coeficientes de las variables originales que son una combinación lineal para obtener la componente principal.

### 3.1.5. PCA-ANOVA

PCA-ANOVA propuesto por [Luciano and Næs, 2009] consiste en usar PCA sobre la matriz de datos originales  $\mathbf{Y}$  de dimensiones  $a*b*n \times p$  (catadores\*productos\*réplicas×atributos), es decir el número de columnas corresponden al número de atributos y el número de filas corresponden a todas las combinaciones de catadores, productos y réplicas. Estos datos se pueden expresar usando un modelo MANOVA formulado en la sección 3.1.3 de la siguiente forma:

$$\mathbf{Y} = \mathbf{1}\mu^T + \boldsymbol{\tau} + \boldsymbol{\beta} + \boldsymbol{\tau}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (16)$$

Lo anterior implica que PCA produce unos componentes que son combinaciones de todos los efectos del modelo anova multivariado, ver (16). Cuando se calculan los factores de PCA, pueden ser interpretados con el uso de gráficas de dispersión y además relacionarlos

con la matriz del diseño de los efectos usando ANOVA de dos factores. El modelo ANOVA de la ecuación (16) puede ser reescrito reemplazando  $\mathbf{Y}$  en la ecuación (15).

$$\mathbf{T} = \mathbf{Y}\mathbf{P} = (\mathbf{1}\mu^T + \tau + \beta + \tau\beta + \varepsilon)\mathbf{P} \quad (17)$$

Se puede trabajar sobre esta matriz empleando ANOVA regularmente, es decir, se obtienen todos los cálculos de los efectos y pruebas de hipótesis sobre la matriz  $\mathbf{T}$ . Debido a que los factores  $\mathbf{T}$  de PCA son ortogonales, se justifica el empleo de ANOVA separado para cada factor.

Este método puede extenderse fácilmente a situaciones donde el efecto de los productos se compone de diferentes factores experimentales, por ejemplo, tomar un producto de varias fábricas que, además del efecto del producto, se quiera estudiar diferentes periodos de almacenamiento. Este ejemplo puede ser modelado como sigue:

$$\mathbf{Y} = \mathbf{1}\mu^T + \tau + \beta + \gamma + \tau\beta + \tau\gamma + \beta\gamma + \varepsilon \quad (18)$$

donde  $\tau$  y  $\beta$  son los dos factores experimentales del producto y  $\gamma$  es el factor del catador. PCA se aplica a  $\mathbf{Y}$  de la misma forma y ANOVA se aplica a los factores  $\mathbf{T}$  de PCA simplemente incorporando un factor en el modelo. Todos los cálculos de los efectos y la pruebas de hipótesis se realizan de la manera usual.

### 3.1.6. Corrección de escala Brockhoff

Sea  $Y_{cpr}$  la réplica  $r$ -ésima de una evaluación sensorial realizada por el catador  $c$  en el producto  $p$ , donde  $r = 1, \dots, R$ ,  $c = 1, \dots, C$  y  $p = 1, \dots, P$ . El modelo ANOVA estándar con efectos principales del producto y el catador, junto con el efecto de la interacción producto $\times$ catador puede ser escrito como:

$$\begin{aligned} Y_{cpr} &= \alpha_c + v_p + \gamma_{cp} + \varepsilon_{cpr}, \\ \varepsilon_{cpr} &\sim N(0, \sigma^2) \quad \text{e independiente} \end{aligned} \quad (19)$$

La aplicación de este modelo tiene el propósito de obtener información del producto ( $v_p$ ) pero teniendo en cuenta posibles diferencias entre los catadores. La figura 3.1 ilustra las diferencias que se pueden tener en cuenta en una escala univariada. Los tipos de diferencias que se tienen en cuenta entre los catadores incluyen: 1) nivel: la diferencia de nivel entre un par de evaluadores se refiere a la percepción subjetiva de la intensidad del atributo en la muestra, por parte de cierto catador. 2) escala: la diferencia de escala se refiere a cómo se perciben y ponderan los incrementos y reducciones de la intensidad del atributo en la muestra. 3) desacuerdo: comprende el error producido por uno o más catadores que evalúan uno o varios productos de forma opuesta a otros catadores. 4) variabilidad: el error de variabilidad se refiere a la repetitividad y precisión que tiene un catador al evaluar un mismo producto en diferentes sesiones de catación. Las diferencias de escala y desacuerdo constituyen, juntas, la interacción catador $\times$ producto.

### Modelo básico de evaluador (catador)

El modelo ANOVA dado en 19 considera las diferencias de tipo 1, 2 y 3 de la figura 3.1. Las diferencias de nivel de los catadores son los efectos principales ( $\alpha_c$ ). Las diferencias de escala y desacuerdo se muestran como el efecto de interacción ( $\gamma_{cp}$ ). Las diferencias de variabilidad no se tienen en cuenta en 19 debido a que una de las suposiciones de los

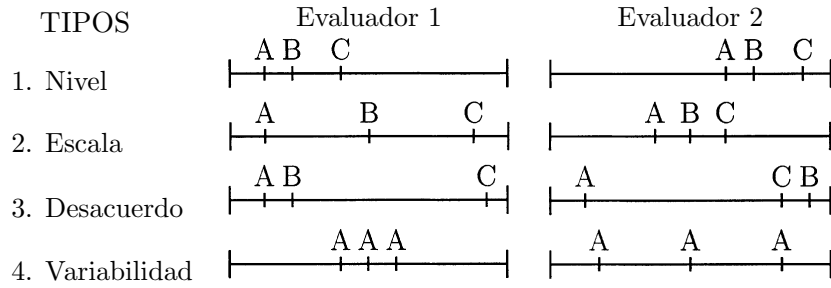


FIGURA 3.1. Diferencias entre evaluadores

modelos ANOVA es la homogeneidad de la varianza. Pero es fácil especificar el mismo modelo permitiendo diferente variación de los evaluadores:

$$Y_{cpr} = \alpha_c + v_p + \gamma_{cp} + \varepsilon_{cpr}, \quad (20)$$

$$\varepsilon_{cpr} \sim N(0, \sigma_{c0}^2) \text{ e independiente}$$

De esta forma, este modelo tiene en cuenta las cuatro diferencias mencionadas entre evaluadores.

Nótese que la única diferencia entre (1) y (2) es el subíndice en la varianza del error  $\sigma_{c0}^2$ . Siguiendo el principio de máxima verosimilitud para probar los efectos se puede sugerir una prueba de desacuerdo: se debe especificar un modelo que incluya cada tipo de diferencias entre evaluadores excepto para el desacuerdo. Este es el modelo básico de evaluador:

$$Y_{cpr} = \alpha_c + \beta_c v_p + \varepsilon_{cpr}, \quad (21)$$

$$\varepsilon_{cpr} \sim N(0, \sigma_c^2) \text{ e independiente}$$

Para explicar como surge este modelo, se tiene en cuenta el modelo (20) como referencia. Para construir un modelo que incluya las diferencias de evaluadores tipo 1, 2 y 4 sin el tipo 3, ver figura (3.1). La inclusión de los parámetros  $\alpha_c$  y  $\sigma_c^2$  asegura la presencia del tipo 1 y 4 de los efectos. El parámetro  $v_p$  (efecto principal del producto) debe estar ahí también puesto que se desea expresar las diferencias de los productos en el modelo. La única cosa que no se usa del modelo (20) es el término de interacción  $\gamma_{cp}$ . Primero, nótese que el efecto de interacción evaluador  $\times$  producto puede verse como la suma de los efectos de escala y efectos no lineales:

$$\text{Interacción} = \text{Escala} + \text{Desacuerdo}$$

En lugar de un término de interacción general  $\gamma_{cp}$  se debe especificar un término que exprese solo los efectos de escala. Las diferencias de escala pueden expresarse especificando constantes de escala individuales  $\tilde{\beta}_c$ . Interpretando el efecto principal de producto  $v_p$  en (20) como el valor real de la propiedad sensorial en cuestión (en una escala común) el efecto de escala está en el producto  $\tilde{\beta}_c v_p$ . Esto debe sustituirse por  $\gamma_{cp}$  en (20) lo que conduce a:

$$Y_{cpr} = \alpha_c + v_p + \tilde{\beta}_c v_p + \varepsilon_{cpr}$$

$$= \alpha_c + (1 + \tilde{\beta}_c)v + \varepsilon_{cpr}$$

renombrando  $\beta_c = (1 + \tilde{\beta}_c)$ , se obtiene el modelo básico de evaluador (21). Esto muestra que aunque a primera vista, los efectos principales del producto hacen falta en el modelo del catador, están presentes como una parte del efecto de escala.

La sensibilidad está definida como la relación *señal a ruido* al cuadrado [Brockhoff, 1998]:

$$Sensibilidad = \frac{\beta_c^2}{\sigma_c^2} \quad (22)$$

Comparando el ajuste de los datos del modelo (21) con el modelo (20) produce una prueba de desacuerdo. De hecho, esta comparación equivale a una comparación de la varianza individual estimada de los dos modelos:  $\hat{\sigma}_{c0}^2$  y  $\hat{\sigma}_c^2$ . El primero es la variación dentro del producto para cada individuo mientras que el último también incluye efectos de desacuerdo (si existen). Entonces se puede obtener una medida individual de desacuerdo comparando los niveles individuales:

$$Desacuerdo_c = \sqrt{\sigma_c^2 - \sigma_{c0}^2} \quad (23)$$

Usando el modelo del evaluador como referencia es posible probar efectos de escala ajustando a una versión restringida sin efectos de escala. También es posible probar diferencias de sensibilidad usando un enfoque similar. Y finalmente, aplicando el modelo del evaluador para cada atributo se obtienen estimadores para escalas individuales, variabilidad, desacuerdo y sensibilidad para cada atributo.

### El modelo generalizado de evaluador

En las situaciones reales, la estructura del diseño de experimentos tiene una complejidad tal que el modelo básico de evaluador sería muy simple. La estructura compleja puede incluir efectos cruzados y enlazados, efectos de orden, de seguimiento, de sesión y efectos de bloque, etc.

El modelo de evaluador general se convierte en

$$Y_{cpr} = \alpha_c + \beta_c(v_p + \delta_e) + \varepsilon_{cpr}, \varepsilon_{cpr} \sim N(0, \sigma_c^2) \quad \text{e independiente} \quad (24)$$

Este es entonces el modelo que permite interacciones multiplicativas entre catadores y cualquier otra cosa, por ejemplo, la temperatura, pero solo a través de las mismas constantes de escalas individuales. Esto se puede explicar así: no importa que pueda cambiar la percepción de los productos, tendrá que pasar por el uso individual de la escala antes de convertirse en una evaluación de la muestra. Desde este punto de vista esta generalización no cambia nada de las pruebas de significancia estadística para las diferencias individuales con las estimaciones individuales de la escala, variabilidad, desacuerdo y sensibilidad.

La versión general del modelo de referencia (20) puede ser escrito como:

$$Y_{cptr} = \alpha_c + v_p + \delta_t + \gamma_{cp} + \eta_{ct} + \varepsilon_{cptr}, \quad (25)$$

$$\varepsilon_{cptr} \sim N(0, \sigma_{c0}^2) \quad \text{e independiente}$$

#### 3.1.7. Corrección para diferentes usos de la escala entre los evaluadores

En [Brockhoff, 2003, Brockhoff and Skovgaard, 1994] se propone un modelo ANOVA denominado modelo de evaluador, que toma en cuenta, además de las diferencias entre los

productos, diferencias en el uso de la escala y diferentes varianzas individuales de los evaluadores. Esta metodología será empleada para corregir las diferencias individuales entre los evaluadores para análisis posteriores. Este modelo puede ser escrito como se explica en la sección 3.1.6 y se indica en la ecuación 26 :

$$Y_{cij}^k = \alpha_c^k + \beta_c v_i^k + e_{cij}^k \quad \text{donde} \quad e_{cij}^k \sim N(0, \sigma_{ck}^2) \quad (26)$$

Para corregir las diferencias individuales se consideran dos opciones. Una es la estandarización tradicional de los datos usando los valores estimados de  $\hat{\alpha}_c^k$  y  $\hat{\beta}_c^k$ , denominado escalado *multiplicativo*:

$$S_{cij}^k = (y_{cij}^k - \hat{\alpha}_c^k) / \hat{\beta}_c^k \quad (27)$$

donde  $S_{cij}^k$  son los datos corregidos usando el escalado *multiplicativo*. Sin embargo, esta transformación no tiene en cuenta el hecho de que las constantes de escala  $\beta_c^k$  proviene de un procedimiento iterativo donde los individuos son también ponderados de acuerdo a su *sensibilidad*, conduciendo posiblemente a resultados sesgados. Para superar este inconveniente, también se considera otro método de escalamiento, llamado *aditivo*:

$$D_{cij}^k = y_{cij}^k - (\hat{\beta}_c^k - \bar{\beta}^k) v_i^k \quad (28)$$

donde  $\bar{\beta}^k$  es el promedio de los valores  $\hat{\beta}_c^k$  estimados y  $D_{cij}^k$  son los datos corregidos usando el escalado *aditivo*. Esta transformación solo permite remover de los datos la parte de interacción,  $\beta_c^k v_i^k$  [Romano et al., 2008].

### 3.1.8. Análisis de Procrustes

Cada matriz correspondiente al evaluador,  $\mathbf{Y}_k$ , consiste de N filas con puntajes para  $d$  atributos. Esta matriz contiene elementos  $\mathbf{Y}_{ik}^l$ , donde  $i$  es el índice sobre los  $P$  productos,  $l = 1, \dots, d$ , es el número de atributos y  $k = 1, \dots, c$  es el número de evaluadores.

El objetivo del análisis de Procrustes es tratar de conseguir que los mismos objetos estén tan cerca como sea posible reduciendo la influencia del catador mediante ajustes de escala y nivel, rotando y reflejando las diferentes configuraciones (arreglo de puntos sobre un hiperplano M-dimensional) si es necesario. La suposición más importante es que las distancias entre los N objetos para un catador no cambien durante estas transformaciones. Aunque las configuraciones pueden ser contraídas o encogidas, las distancias relativas entre los objetos permanecen invariantes.

Las distancias entre los objetos reflejan las relaciones entre los objetos. Los objetos cercanos son similares, mientras los objetos alejados son diferentes. La razón de mantener las distancias invariantes es que en este proceso de ajuste, las relaciones entre los N objetos de un evaluador no deben cambiar. Los objetos similares deben continuar siendo similares, y los objetos diferentes deben seguir siendo diferentes.

Esta técnica entrega una transformación de los datos originales que consiste en transformar la matriz de cada catador resolviendo las diferencias producidas por el efecto de nivel, el efecto de interpretación y el efecto de rango, pero con la restricción de conservar las proporciones entre los productos de cada catador. En otras palabras, la transformación de



Procrustes permite mejorar la concordancia entre los catadores y obtener así una medida promedio más confiable en la etapa de integración de variables.

Matemáticamente el proceso de ajuste es expresado minimizando las distancias entre los mismos objetos para diferentes catadores, bajo la condición de que las distancias entre los objetos de un mismo catador no cambien. Estas distancias pueden ser expresadas como las diferencias entre las matrices individuales:

$$d(\mathcal{T}(\mathbf{X}_k), \mathcal{T}(\mathbf{X}_{k'})) = \sum_{k < l}^K \|\mathcal{T}(\mathbf{X}_k) - \mathcal{T}(\mathbf{X}_{k'})\| \quad (29)$$

donde  $\mathcal{T}(\mathbf{X})$  se refiere a cierta transformación  $\mathcal{T}$  de la matriz  $\mathbf{X}$ . Dicha transformación debe conservar las distancias relativas entre las observaciones de  $\mathbf{X}$  y consiste en minimizar

$$\sum_{k=1}^K \|\mathcal{T}(\mathbf{X}_k) - \mathbf{Y}\| \quad (30)$$

donde

$$\mathbf{Y} = K^{-1} \sum_{k=1}^K \mathcal{T}(\mathbf{X}_k) \quad (31)$$

es la media de las matrices de datos individuales transformadas  $\mathcal{T}(\mathbf{X}_k)$ . Las transformaciones aplicadas en el análisis de Procrustes son traslaciones, rotaciones y escalamiento; esta transformación puede ser expresada así

$$\mathcal{T}(\mathbf{X}_k) = \rho_k \mathbf{X}_k \mathbf{H}_k + \mathbf{T}_k \quad (32)$$

donde  $\rho_k$  es el factor de escalamiento,  $\mathbf{H}_k$  es la matriz de rotación y  $\mathbf{T}_k$  es la traslación. La traslación se puede conseguir centrando cada una de las matrices  $\mathbf{X}$  que se quieren transformar. [Dijksterhuis, 1996]

### 3.1.9. Combinación de clasificadores

La estrategia de combinación de clasificadores se emplea en circunstancias donde existen varias fuentes de adquisición de datos de la misma naturaleza, o bien, varios conjuntos de datos de diferente naturaleza obtenidos de los mismos objetos. Normalmente es recomendable emplear reglas de clasificación sencillas, tales como el clasificador del vecino más cercano y análisis discriminante lineal, de tal forma que se entrene un clasificador para cada conjunto de datos y posteriormente se obtenga una decisión que puede ser tomada en consenso o basada en la mayor probabilidad a posteriori de los clasificadores individuales. En (Granitto et al. 2008) se propone el empleo de combinación de clasificadores para aprovechar las habilidades individuales de los catadores. Entre las reglas de combinación más sencillas y estables se encuentra la denominada regla de veto. Mediante una analogía con un comité de expertos, esta regla se puede explicar como la imposición de la decisión tomada por el experto que esté más seguro de su opinión; es decir, clasificar cada observación de acuerdo con la probabilidad a posteriori más alta entre todos los clasificadores individuales.

### 3.2. Cromatografía

La cromatografía es una técnica de separación de compuestos químicos de una mezcla compleja permitiendo su identificación y determinación de la cantidad de dicho componente en la muestra de interés. Las técnicas cromatográficas son variadas, sin embargo, en todas ellas existe una fase móvil (consiste en un gas o líquido) la cual conduce a los componentes químicos a través de una fase estacionaria que es un sólido o un líquido fijo en una columna, con la cual dichos compuestos interactúan a distintas velocidades y se van separando hasta ingresar a un detector, el cual genera una señal formada por picos correspondientes al tipo de compuesto y la cantidad de éste presente en la muestra de análisis. En particular, la cromatografía de gases (GC) es una técnica usada para separar y analizar compuestos orgánicos que son volátiles. El método se basa en la inyección de una muestra preparada previamente en un puerto inyector del cromatógrafo de gases el cual se encuentra a temperaturas mayores a las de ebullición de los componentes, ayudando a la evaporación de la mezcla y permitiendo el transporte de la misma, con ayuda de la fase móvil, a la columna compuesta de una fase estacionaria donde los compuestos son separados. Posteriormente cada uno de los componentes ingresa al detector en un tiempo característico para cada compuesto particular, conocido como tiempo de retención. Dicho detector envía una señal que es registrada como un pico en el software del computador y cuya área es proporcional al número de moléculas presentes en el compuesto [P.W.Scott, 1998] (ver Figura 3.2).

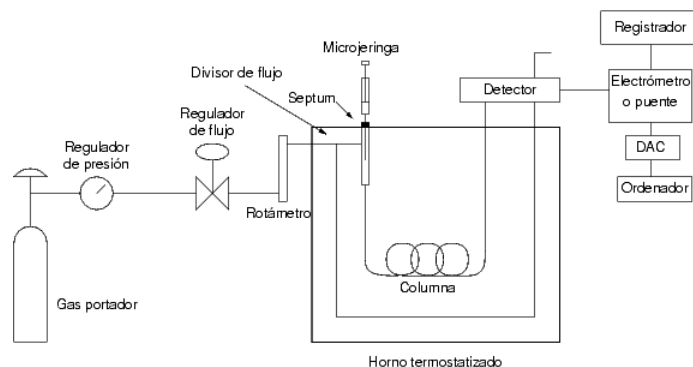


FIGURA 3.2. Equipo de cromatografía de gases

Los perfiles de cromatografía de gases (GC), conocidos como *fingerprinting*, es decir la identificación de una muestra análoga a una huella digital, son usados generalmente para distinguir una muestra de otra, realizar control de calidad de cierto producto o caracterizar una muestra basada en sus compuestos químicos. Un proyecto a gran escala de *fingerprinting*, como por ejemplo el control de calidad de un proceso en cadena o por etapas, usando

muestras de GC en un periodo de tiempo largo, puede requerir una comparación sistemática de cientos y miles de perfiles de cromatografía adquiridos en diferentes días, meses, años, posiblemente de diferentes instrumentos GC.

Los datos de cromatografía exhiben incertidumbre sobre el eje del tiempo debido a variaciones inevitables en los parámetros del instrumento, por ejemplo, cambios en la temperatura y el flujo entre corrida y corrida, efectos de las diferencias entre muestra y muestra, y desgaste a través del tiempo de las partes internas del cromatógrafo, todos conducen a variaciones en el tiempo de retención para un análisis dado sobre una serie de muchas corridas de cromatografía.

Las técnicas de quimiometría son sensibles a la precisión del tiempo de retención. Por ejemplo, dos cromatogramas con un desplazamiento de uno con respecto al otro, aunque sean idénticos en cualquier otro aspecto, como cantidad, forma, área de los picos, tendería a ser interpretado incorrectamente por las técnicas de análisis quimiométrico como si fueran químicamente diferentes. Para encontrar acertadamente diferencias en la composición química entre muestras con base en su cromatografía, debe haber una correspondencia directa entre las variables de los dos objetos a ser comparados. Puede ocurrir en ciertas ocasiones que un desplazamiento en el cual cada punto del cromatograma está corrido la misma cantidad de puntos con respecto al cromatograma de referencia, aunque es muy raro, se debería a un problema de "injection-timing". Más comunmente, se encuentra que hay un estrechamiento o estiramiento sobre el eje del tiempo de retención con respecto a otro, esto ocurre cuando un parámetro es diferente globalmente entre dos cromatogramas, tales como velocidad del flujo, temperatura, o la temperatura programada. Un tercer modo de desplazamiento inducido por el instrumento es debido a fluctuaciones más rápidas de estos mismos parámetros durante una corrida de cromatografía lo cual es irreproducible de corrida a corrida.

Mientras que las fuentes de desplazamiento con respecto al instrumento han sido reducidas para GC a través del desarrollo de sistemas de control electrónico y las fuentes de desplazamiento relacionadas químicamente pueden minimizarse ajustando las condiciones de la separación cromatográfica y reemplazando las partes antes de que ocurra un desgaste significativo, generalmente persiste una cantidad sutil de variación en el tiempo de retención en conjuntos grandes de cromatografías tomados en largos periodos de tiempo.

### **3.2.1. Alineamiento dinámico en el tiempo (DTW)**

El método DTW (*Dynamic time warping*), es un algoritmo de alineamiento temporal no lineal, que se usa comúnmente en la normalización en longitud y en la estimación de la disimilitud entre dos señales. En algunas ocasiones las señales o vectores representativos son de longitudes diferentes y estas medidas no pueden aplicarse directamente. Entonces, en estos casos, debe hacerse una normalización temporal previa, que consiste en ajustar las longitudes de las señales a un mismo valor, para que los vectores puedan ser comparados a través de alguna de las medidas mencionadas. La normalización puede llevarse a cabo con métodos lineales y no lineales. En la normalización temporal lineal se escala el vector de

menor longitud para que los dos vectores a comparar tengan la misma longitud. Así, para calcular la disimilitud entre dos señales se deben repetir o interpolar muestras de la señal de longitud menor hasta que coincida con la longitud de la otra señal. De esta manera, se puede interpretar los vectores como elementos de un espacio vectorial y se puede aplicar medidas como la distancia euclídea.

El método de alineamiento temporal no lineal DTW, mejora las deficiencias de la normalización temporal lineal y se basa en ajustar únicamente la longitud de los tramos de las señales que resulte más conveniente desde el punto de vista de su comparación, sin alterar el resto. Todo esto se logra en un esquema de programación dinámica. Por esta razón el algoritmo DTW es de uso frecuente en la clasificación de muestras de longitudes diferentes. La programación dinámica se puede implementar con una multitud de variaciones de acuerdo a su aplicación.

### Algoritmo DTW

El valor de la disimilitud se obtiene en el último elemento de una matriz, denominada matriz de programación dinámica  $\mathbf{G}$ , que se denota con  $G[n_x - 1, n_y - 1]$ , donde  $n_x$  y  $n_y$  son las longitudes de los vectores  $\mathbf{x}$  e  $\mathbf{y}$  a comparar. Este valor debe normalizarse respecto a la longitud del camino de alineamiento para evitar el efecto de la longitud. La matriz  $\mathbf{G}$  tiene tantas filas como la longitud del primer vector y tantas columnas como la longitud del segundo vector, como se muestra en la Figura 3.3. Esta matriz se construye de forma que cada uno de sus elementos o nodos, represente el costo acumulado del alineamiento entre los vectores hasta llegar a ese punto.

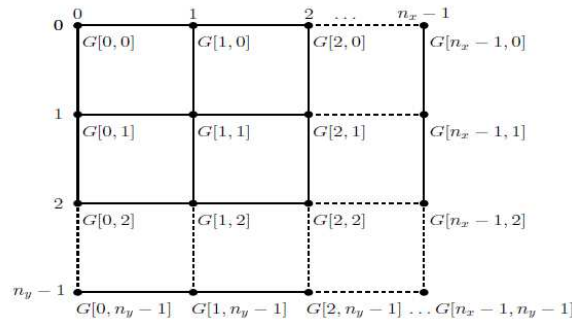


FIGURA 3.3. Matriz de programación dinámica  $\mathbf{G}$  asociada a los vectores  $x$  e  $y$

La matriz de programación dinámica es el plano discreto en donde se traza el camino de alineamiento. El objetivo del algoritmo consiste en encontrar el camino de alineamiento de costo mínimo entre los dos vectores. Este camino se definirá como el conjunto ordenado de nodos que se encuentran desde el nodo inicial  $(i_1, j_1)$  hasta el nodo final  $(i_N, j_N)$ , de la forma:  $(i_1, j_1), (i_2, j_2) \dots (i_N, j_N)$ .

Para establecer el costo del camino es necesario definir el costo debido a la transición entre cada par de nodos. La función de costo asociada a la  $k$ -ésima transición se denota

como  $d[i_{k-1}, j_{k-1} | (i_k, j_k)]$ . El costo debe ser un valor positivo y se puede medir a través de la norma  $L_1$  ó  $L_2$ .

El costo asociado a un camino completo puede calcularse con la suma del costo de cada una de las transiciones:

$$D = \sum_{k=1}^K d[(i_{k-1}, j_{k-1}) | (i_k, j_k)] \quad (33)$$

donde  $K$  representa la cantidad de nodos del camino.

Adicionalmente, es necesario aplicar un factor de ponderación a las transiciones. Este factor se relaciona con el conjunto de transiciones o producciones y representa la longitud correspondiente a la transición. El costo acumulado ponderado es de la forma:

$$D = \sum_{k=1}^K d[(i_{k-1}, j_{k-1}) | (i_k, j_k)] w(a_k, b_k) \quad (34)$$

donde  $w(a_k, b_k)$  representa el factor de ponderación debido a la transición de  $(i_k, j_k)$  a  $(i_{k-1}, j_{k-1})$ .

Finalmente, se realiza una normalización con respecto de la longitud, así:

$$D = \frac{\sum_{k=1}^K d[(i_{k-1}, j_{k-1}) | (i_k, j_k)] w(a_k, b_k)}{\sum_{k=1}^K w(a_k, b_k)} \quad (35)$$

Haciendo  $N = \sum_{k=1}^K w(a_k, b_k)$ , se tiene que

$$D = \frac{G[n_x - 1, n_y - 1]}{N} \quad (36)$$

La expresión recursiva para la obtención de cada elemento de la matriz  $G$  es:

$$G[i, j] = \min_{a,b} \sum_{k=1}^K d[(i_{k-1}, j_{k-1}) | (i_k, j_k)] w(a_k, b_k) \quad (37)$$

donde  $G[i - a, j - b]$  representa el costo mínimo acumulado considerando el conjunto de producciones  $P$ .

### 3.2.2. Alineamiento de correlación optimizado (COW)

Es un método de preprocesamiento de señales que opera a tramos, sobre una única muestra a la vez. Tiene por objetivo alinear el vector de datos de la señal de muestra con respecto a un vector de referencia permitiendo cambios limitados en las longitudes de los segmentos en el vector de muestra. El máximo incremento o decremento en la longitud en un segmento de la muestra se controla mediante un parámetro denominado dilatación  $t$ . Cuando difieren el

número de puntos en el eje del tiempo entre la muestra correspondiente y el segmento de referencia, la primera se interpola linealmente con el fin de crear un segmento de igual longitud.

En COW, las diferentes longitudes de segmento en el vector de muestra se seleccionan (o se desplazan los bordes; se "alinean") de modo que se optimice la correlación total entre la muestra y la referencia. El problema se resuelve descomponiendo el problema global en una optimización de la correlación por segmentos mediante un algoritmo de programación dinámica. El espacio solución de esta optimización se define mediante dos parámetros: el número de bordes de segmentos  $I + 1$  y la longitud del área de dilatación  $t$ . Es convencional fijar las fronteras inicial y final de modo que los primeros y últimos puntos en los vectores de muestra y de referencia se obligan a ajustarse. Los detalles más particulares de la implementación se encuentran [Nielsen et al., 1998]. El procedimiento de optimización de los parámetros  $I$  y  $t$  presentados arriba se expone en [Skov et al., 2006].

### 3.2.3. Algoritmo de ajuste de picos

En el enfoque de alineamiento propuesto en [Johnson et al., 2003] para ajustar picos con respecto a su tiempo de retención (TR), los picos de las cromatografías son por sí mismos las características en común entre la muestra y la referencia. Sin embargo, el proceso de alineamiento requiere picos correspondientes de cromatografía a cromatografía que no estén desplazado en el tiempo de retención por más de cierta distancia típica entre los picos adyacentes. Si este requisito no se cumple, es necesario aplicar previamente un procedimiento de ajuste de picos. El pico del cromatograma muestra es ajustado al pico del cromatograma de referencia que tiene el TR más cercano a éste, dentro de una distancia umbral dada en el eje de TR. Las regiones entre los picos en el cromatograma muestra son alargadas o encogidas interpolando más o menos puntos de tal forma que se fuerzan los picos en la muestra para que ocurran al mismo tiempo de retención que los picos de la referencia.

El cromatograma de referencia debe ser escogido cuidadosamente. Es importante que los picos de la referencia esté ubicados tan cerca como sea posible de los centros de distribución de las posiciones de todos los picos de la base de datos que serán alineados. También es deseable que la referencia contenga, en la medida de lo posible, todos los picos contenidos por el resto de cromatogramas en el conjunto de datos que serán alineados.

En el primer paso del proceso de alineamiento se identifican los picos cromatográficos en ambas cromatografías, la muestra y la referencia, y se genera una lista de sus TR. Los picos se ubican automáticamente encontrando cruces por cero de la estimación de la primera derivada del cromatograma. El algoritmo opera en un perfil de cromatografía a la vez y las procesa paso a paso a través de cada punto en el cromatograma. Teóricamente, la diferencia máxima observada en el ruido de línea base podría aproximarse a cuatro veces la desviación estándar del ruido. Cuando esta diferencia se incrementa más allá de un umbral cinco veces mayor que la desviación estándar del ruido de línea base, el algoritmo ha detectado el borde inicial del pico de cromatografía, y empieza buscando un cruce por cero en la estimación de la derivada. Este pico es guardado en una lista de tiempos de retención de picos.

La ubicación de los picos con relaciones señal a ruido más pequeñas es problemática debido a la posibilidad de que sea un pico que está por debajo del umbral de búsqueda de picos.

En seguida, los cromatogramas de muestra y de referencia son comparados paso a paso a través de cada uno de los picos en el cromatograma de referencia y encontrando el pico del cromatograma muestra más cercano que se ajuste en el TR. Si el ajuste más cercano está dentro del ancho de la ventana del pico seleccionado para ajustar, después se asume que el pico no está presente en la muestra y no es usado en el alineamiento de TR, de esta forma se permite alinear cromatogramas con diferente cantidad de picos. El eje del tiempo de retención de la muestra de cromatografía es ajustada a través de interpolación tal que los picos ajustados tengan el mismo tiempo de retención. Esto se consigue interpolando más o menos puntos de datos entre los picos de cromatografía tal que se expanda o comprima el eje del tiempo de retención. Una medida común para determinar la calidad del alineamiento de perfiles de cromatografía similares es el coeficiente de correlación entre ellos. Dos cromatogramas que son idénticos tienen un coeficiente de correlación igual a uno. Detalles adicionales así como un código fuente de la implementación están disponibles<sup>1</sup> en [Pierce et al., 2005]

### 3.3. Integración de datos cualitativos y cuantitativos

#### 3.3.1. Regresión por mínimos cuadrados parciales (PLSR)

PLSR es un método para relacionar dos matrices de datos,  $\mathbf{X}$  y  $\mathbf{Y}$ , mediante un modelo multivariado lineal, pero va más allá de la regresión tradicional en el hecho de que modela también la estructura de  $\mathbf{X}$  y  $\mathbf{Y}$ . PLSR deriva su utilidad de su habilidad para analizar datos con numerosas variables, ruidosas, colineales, e incluso incompletas en ambas matrices  $\mathbf{X}$  y  $\mathbf{Y}$ . PLSR tiene la deseable propiedad de que la precisión de los parámetros del modelo mejora con el incremento del número de variables relevantes y observaciones. PLSR es un método de modelamiento multivariado cuantitativo, con posibilidades inferenciales, similar a regresión múltiple, prueba-t y ANOVA.

El modelo lineal PLSR encuentra unas pocas “nuevas” variables, que son estimaciones de las variables latentes o sus rotaciones. Estas nuevas variables son llamadas puntajes de X y son denotadas por  $t_a$  ( $a = 1, 2, \dots, A$ ). Los puntajes de X son predictoras de  $\mathbf{Y}$  y también modelan  $\mathbf{X}$ , i.e., ambas  $\mathbf{Y}$  y  $\mathbf{X}$  se asumen, al menos en parte, modeladas por las mismas variables latentes.

Los A puntajes de X proporcionan una alta reducción de dimensión y son ortogonales. Dichos puntajes son estimados como combinaciones lineales de las variables originales  $x_k$  con los coeficientes, “pesos”,  $w_{ka}^*$  ( $a = 1, 2, \dots, A$ ). Abajo se muestran las fórmulas en ambas formas, como elementos o matrices:

$$t_{ia} = \sum_k W_{ka}^* X_{ik} \quad (\mathbf{T} = \mathbf{XW}^*) \quad (38)$$

<sup>1</sup><http://synoveclab.chem.washington.edu/Downloads.htm>

Los puntajes de  $X$  ( $t_a$ 's) tienen las siguientes propiedades: (a) Son, multiplicados por las cargas  $p_{ak}$ , buenos "resúmenes" de  $X$ , tal que los residuos de  $X$ ,  $e_{ik}$ , en la ecuación (39) son "pequeños":

$$X_{ik} = \sum_a t_{ia} p_{ak} + e_{ik} \quad (\mathbf{X} = \mathbf{TP}^T + \mathbf{E}) \quad (39)$$

Con  $Y$  multivariado (cuando  $M > 1$ ), los correspondientes "puntajes de  $Y$ " ( $u_a$ ) son, multiplicados por los pesos  $c_{am}$ , buenos "resúmenes" de  $Y$ , tal que los residuos,  $g_{im}$ , en la ecuación (40) son "pequeños":

$$y_{im} = \sum_a u_{ia} c_{am} + g_{im} \quad (\mathbf{Y} = \mathbf{UC}^T + \mathbf{G}) \quad (40)$$

(b) los puntajes de  $X$  son buenas predictoras de  $Y$ , i.e.:

$$y_{im} = \sum_a c_{ma} t_{ia} + f_{im} \quad (\mathbf{Y} = \mathbf{TC}^T + \mathbf{F}) \quad (41)$$

Los residuos de  $Y$ ,  $f_{im}$  expresan la desviación entre las respuestas observadas y las modeladas, y comprende los elementos de la matriz de residuos,  $\mathbf{F}$ .

Debido a que las ecuaciones (38), (41) pueden ser reescritas para que sean vistas como un modelo de regresión múltiple:

$$y_{im} = \sum_a c_{ma} \sum_k w_{ka}^* x_{ik} + f_{im} = \sum_k b_{mk} x_{ik} + f_{im} \\ (\mathbf{Y} = \mathbf{XW}^* \mathbf{C}^T + \mathbf{F} = \mathbf{XB} + \mathbf{F}) \quad (42)$$

Los "coeficientes de la regresión PLS",  $b_{mk}$  ( $\mathbf{B}$ ), pueden ser escritos como:

$$b_{mk} = \sum_a c_{ma} w_{ka}^* \quad (\mathbf{B} = \mathbf{W}^* \mathbf{C}^T) \quad (43)$$

Después de que cada componente,  $a$ , de la matriz  $X$  es extraído sustrayendo  $t_{ia} p_{ka}$  de  $x_{ik}$  ( $t_a p_a^T$  de  $X$ ). Esto hace que el modelo PLSR pueda ser expresado alternativamente en pesos  $w_a$  refiriéndose a los residuos después de la dimensión previa,  $\mathbf{E}_{a-1}$  en lugar de relacionar con las mismas variables  $X$ . Así, en lugar de la ecuación (38), se puede escribir:

$$t_{ia} = \sum_k w_{ka} e_{ik,a-1} \quad (\mathbf{t}_a = \mathbf{E}_{a-1} \mathbf{W}_a) \quad (44a)$$

$$e_{ik,a-1} = e_{ik,a-2} - t_{i,a-1} p_{a-1,k} \quad (\mathbf{E}_{a-1} = \mathbf{E}_{a-2} - \mathbf{t}_{a-1} \mathbf{p}_{a-1}^T) \quad (44b)$$

$$e_{ik,0} = X_{ik} \quad (\mathbf{E}_0 = \mathbf{X}) \quad (44c)$$

Sin embargo, los pesos,  $w$ , pueden ser transformados en  $w^*$ , que se relacionan directamente con  $X$ , dado en la ecuación (38). La relación entre los dos está dada por:

$$\mathbf{W}^* = \mathbf{W}(\mathbf{P}^T \mathbf{W})^{-1} \quad (45)$$



La matriz  $\mathbf{Y}$  también puede ser reducida sustrayendo  $\mathbf{t}_a \mathbf{c}_a^T$ , pero esto no es necesario; los resultados son equivalentes con o sin reducir  $\mathbf{Y}$ .

Del algoritmo PLSR, uno puede ver que el primer vector de pesos ( $\mathbf{w}_1$ ) es el primer vector propio de la matriz combinada varianza-varianza,  $\mathbf{X}^T \mathbf{Y} \mathbf{Y}^T \mathbf{X}$ , y los siguientes vectores de pesos (componente  $a$ ) son los vectores propios de las versiones reducidas de la misma matriz, i.e.,  $\mathbf{Z}_a^T \mathbf{Y} \mathbf{Y}^T \mathbf{Z}_a$ , donde  $\mathbf{Z}_a = \mathbf{Z}_{a-1} - \mathbf{T}_{a-1} \mathbf{P}_{a-1}^T$ . Similarmente, el primer vector de los puntajes de  $\mathbf{X}$  ( $\mathbf{t}_1$ ) es el vector propio de  $\mathbf{X} \mathbf{X}^T \mathbf{Y} \mathbf{Y}^T$ , y después los vectores puntajes de  $\mathbf{X}$  ( $\mathbf{t}_a$ ) son los vectores propios de  $\mathbf{Z}_a \mathbf{Z}_a^T \mathbf{Y} \mathbf{Y}^T$ .

Estas relaciones entre vectores propios también muestra que los vectores  $\mathbf{w}_a$  forman un conjunto ortonormal, y que los vectores  $\mathbf{t}_a$  son ortogonales entre sí. Los vectores de cargas  $\mathbf{p}_a$  no son ortogonales entre sí, y tampoco lo son los puntajes de  $\mathbf{Y}$ ,  $\mathbf{u}_a$ . Los vectores  $\mathbf{u}$  y  $\mathbf{p}$  son ortogonales a los vectores  $\mathbf{t}$  y los  $\mathbf{w}$ , respectivamente, de los componentes anteriores, i.e.,  $\mathbf{u}_b^T \mathbf{t}_a = 0$  y  $\mathbf{p}_b^T \mathbf{w}_a = 0$ , si  $b > a$ . Además,  $\mathbf{w}_a^T \mathbf{p}_a = 1$ .

### 3.3.2. Análisis de correlación canónica (CCA)

Los datos para el análisis de correlación canónica se dividen en dos conjuntos de variables, de forma que cada conjunto tiene un significado teórico: (1) variables independientes,  $\mathbf{X}$  y (2) variables dependientes,  $\mathbf{Y}$  [Hair et al., 1998]. El análisis de correlación canónica busca encontrar la dependencia que existe entre la combinación lineal de un conjunto de variables y la combinación lineal de otro. Estas combinaciones lineales son también denominadas *variables canónicas*. El procedimiento se orienta a la maximización del *coeficiente de correlación*, con el objetivo de concentrar la relación de alta dimensionalidad entre unos cuantos pares de variables canónicas [Johnson and Wichern, 2002].

En general, el problema se puede resumir en el hallazgo de dos variables resumen, una de cada conjunto, que tengan correlación máxima. La variable canónica  $\mathbf{x}^*$  es la combinación lineal de las  $p$  primeras variables, y la variable canónica  $\mathbf{y}^*$  es la combinación lineal de las  $q$  siguientes variables

$$\mathbf{x}^* = \mathbf{X} \boldsymbol{\alpha} = \sum_{i=1}^p \alpha_i x_i \quad (46)$$

$$\mathbf{y}^* = \mathbf{Y} \boldsymbol{\beta} = \sum_{j=1}^q \beta_j y_j \quad (47)$$

Se desea encontrar los vectores  $\boldsymbol{\alpha}$  y  $\boldsymbol{\beta}$  para que las nuevas variables  $\mathbf{x}^*$  y  $\mathbf{y}^*$  tengan correlación máxima. Estas variables se denominan variable canónica predictora  $\mathbf{x}^*$  y variable canónica criterio  $\mathbf{y}^*$ . Para tal fin, la función objetivo a maximizar es

$$\rho^2 = \frac{(\boldsymbol{\alpha}^T \mathbf{S}_{xy} \boldsymbol{\beta})^2}{(\boldsymbol{\alpha}^T \mathbf{S}_{xx} \boldsymbol{\alpha})^2 (\boldsymbol{\beta}^T \mathbf{S}_{yy} \boldsymbol{\beta})^2} = \frac{(\boldsymbol{\alpha}^T \mathbf{R}_{xy} \boldsymbol{\beta})^2}{(\boldsymbol{\alpha}^T \mathbf{R}_{xx} \boldsymbol{\alpha})^2 (\boldsymbol{\beta}^T \mathbf{R}_{yy} \boldsymbol{\beta})^2} \quad (48)$$

donde  $\rho$  es el coeficiente de correlación;  $\mathbf{S}_{xx}$  es la matriz de dimensiones  $p \times p$  de covarianza de las  $x$ ;  $\mathbf{S}_{xy}$  es la matriz de dimensiones  $p \times q$  de covarianza entre las  $x$  y las  $y$ ; y  $\mathbf{S}_{yy}$

es la matriz de dimensiones  $q \times q$  de covarianza de las  $y$ . Además,  $\mathbf{R}_{xx}$ ,  $\mathbf{R}_{xy}$  y  $\mathbf{R}_{yy}$  son las matrices de correlación correspondientes. A continuación se definen las matrices  $\mathbf{A}$  y  $\mathbf{B}$  como

$$\mathbf{A}_{p \times p} = \mathbf{R}_{xx}^{-1} \mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} \mathbf{R}_{xy}, \quad \mathbf{A}\boldsymbol{\alpha} = \rho^2 \boldsymbol{\alpha} \quad (49)$$

$$\mathbf{B}_{q \times q} = \mathbf{R}_{xx}^{-1} \mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} \mathbf{R}_{xy}, \quad \mathbf{B}\boldsymbol{\beta} = \rho^2 \boldsymbol{\beta} \quad (50)$$

donde  $\boldsymbol{\alpha}$  es el vector propio estandarizado ligado al valor propio  $\rho^2$  de la matriz cuadrada  $\mathbf{A}$  y  $\boldsymbol{\beta}$  es el vector propio estandarizado ligado al valor propio  $\rho^2$  de la matriz cuadrada  $\mathbf{B}$ . Los valores propios de las matrices  $\mathbf{A}$  y  $\mathbf{B}$ ,  $\rho^2$ , son iguales.

Tomando  $r$  como el  $\min\{p, q\}$ , el proceso para obtener las  $2r$  variables canónicas, i.e.,  $(x_1, x_2, \dots, x_r)$ ,  $(y_1, y_2, \dots, y_r)$ , consiste en obtener los valores y vectores propios de las matrices  $\mathbf{A}$  y  $\mathbf{B}$ . Estas matrices tienen rango igual a  $r$ , y la obtención de los  $r$  vectores propios asociados a los valores propios, forman  $r$  combinaciones lineales de las variables de ambos grupos que tienen correlación máxima [na, 2002]. Estas variables canónicas forman la primera *función canónica* de  $r$  funciones canónicas que explican la relación total entre los dos conjuntos de variables. La primera función canónica se calcula con el fin de obtener la mayor inter-correlación posible entre los dos conjuntos de variables. La segunda función canónica es obtenida posteriormente para representar la máxima relación entre los dos conjuntos de variables que no se ha explicado por el primer par de variables canónicas, esta segunda función y las funciones canónicas sucesivas son ortogonales e independientes de las anteriores funciones canónicas. Las sucesivas funciones canónicas están basadas en la varianza residual de la muestra.

La validez de la relación entre las variables canónicas se refleja en el coeficiente de correlación canónica. Cuando se eleva al cuadrado, el coeficiente de correlación canónica representa la cantidad de varianza de una variable canónica explicada por la otra variable o la cantidad de varianza compartida entre las dos variables canónicas. Se recomienda que sean empleados dos criterios conjuntamente para decidir qué funciones canónicas se deben interpretar. Los criterios son: 1) el nivel de significación estadística de las funciones y 2) la medida de la redundancia para el porcentaje de varianza explicado por los conjuntos de datos. La magnitud de las correlaciones canónicas podría considerarse como un tercer criterio de decisión, no obstante este criterio carece de objetividad y no es recomendable tenerlo en cuenta [Hair et al., 1998].

El **índice de redundancia** de Stewart-Love [Hair et al., 1998] calcula la cantidad de varianza de un conjunto de variables que puede ser explicada por la varianza de otro conjunto. El conjunto de variables predictoras se espera que explique solamente la varianza compartida del valor teórico canónico criterio. Por esta razón, el cálculo del índice de redundancia es un proceso en dos pasos. El primer paso comprende calcular la cantidad de varianza compartida del conjunto de variables dependientes incluida en el valor teórico canónico dependiente. El segundo paso consiste en calcular la cantidad de varianza en el valor teórico criterio que

puede ser explicada por el valor teórico canónico independiente. El paso final es calcular el índice de redundancia, que se calcula multiplicando estos dos componentes.

### 3.4. Estado del arte

El estudio de relaciones entre evaluaciones sensoriales y datos instrumentales ha sido un tema de gran interés desde finales de los años 80 [Næs and Kowalski, 1989, Parolari et al., 1992, Williams et al., 1988]. De hecho, las técnicas de análisis multivariado, quimiometría y reconocimiento de patrones son las más empleadas en este tipo de estudios. En particular, técnicas como regresión múltiple, regresión por mínimos cuadrados parciales (PLSR), correlación canónica (CCA), análisis de Procrustes (GPA) [Chung et al., 2003, Du et al., 2010], análisis de componentes principales (PCA) [Lassoued et al., 2008, Chung et al., 2003, Kallithraka et al., 2001], análisis de variables canónicas (CVA) [Chung et al., 2003, Adhikari et al., 2003], entre muchos otros, han sido ampliamente usados para tratar de identificar los compuestos y reacciones químicas que generan los aromas y sabores en productos alimenticios [Scacco et al., 2007, Arvanitoyannis et al., 2007, Verzera et al., 2008]. En particular, PCA es empleado principalmente para hacer exploración de la estructura de los datos instrumentales y CVA para hacer exploración de la estructura de los datos sensoriales.

En general, el análisis cuantitativo descriptivo ha sido empleado para la obtención de perfiles sensoriales, puesto que debido a su metodología en la evaluación de los atributos sensoriales, dichas evaluaciones son tomadas como medidas continuas, facilitando el análisis mediante técnicas de análisis multivariado [Du et al., 2010, Bro et al., 2008]. Adicionalmente, es una práctica común obtener un consenso entre el panel sensorial, promediando las evaluaciones entre los catadores, con el fin de hacer más robustos los perfiles sensoriales, lo cual es viable cuando el panel está compuesto por diez o más expertos [Jørgensen and Næs, 2008, Aishima, 2004]. No obstante, estudios empleando análisis de Procrustes (GPA) han mostrado que los productos son percibidos de forma diferente por los miembros del panel [Lassoued et al., 2008, Dijksterhuis, 1994]. En [Adhikari et al., 2003], se encontró que el panel fue muy consistente como grupo en sus evaluaciones. Sin embargo, en otros estudios, la consistencia de los catadores fue evaluada para encontrar si la amplia selección de productos tiene algún efecto en la habilidad de los catadores para detectar diferencias y discriminar entre los productos analizados [Brockhoff and Skovgaard, 1994, François et al., 2008, Jørgensen and Næs, 2008].

En [Dijksterhuis, 1994] se clasificaron los métodos estadísticos de relación de variables en simétricos y asimétricos. Los primeros son capaces de predecir un conjunto de datos a partir de otro, mientras que los segundos simplemente describen la relación entre los conjuntos de datos. GPA es un método simétrico que extrae una estructura común de múltiples conjuntos de datos compuestos por variables de diferente naturaleza, e.g., sensorial, instrumental; En GPA se genera una transformación en consenso que ajusta la configuración entre los múltiples conjuntos de datos centrando, rotando y ajustando las escalas de cada conjunto de datos. PLSR es un método asimétrico usado para entender las relaciones entre dos conjuntos de datos haciendo predicción de un conjunto de datos a partir de otro. Este método busca encontrar la mejor solución tal que  $X$  explique la variación de los datos en  $Y$ .

El método PLSR ha mostrado ser conveniente y ha sido ampliamente usado para la identificación de relaciones entre dos conjuntos de datos, generalmente el primero compuesto por una o más variables dependientes y el segundo compuesto por una gran cantidad de variables explicativas. Este método presenta ciertas características de gran conveniencia en el análisis quimiométrico, como por ejemplo poder usar variables provenientes de señales o espectros, tales como cromatografía de gases acoplada a espectrometría de masas (GC-MS) [Du et al., 2010, Verzera et al., 2008, Durante et al., 2006], espectroscopía FT-IR [Jørgensen and Næs, 2008], espectroscopía del infrarrojo cercano [Cozzolino et al., 2006], entre otros, que se caracterizan por generar una gran cantidad de variables colineales. Otra ventaja de dicho método es que está diseñado para manejar conjuntos de datos donde el número de variables es mucho más grande que el número de observaciones [Höskuldsson, 1988]. PLS es empleado frecuentemente para problemas quimiométricos, algunas aplicaciones pueden encontrarse en [Durante et al., 2006, Adhikari et al., 2003, Chung et al., 2003, François et al., 2008, Jørgensen and Næs, 2008]. Particularmente, en [Aznar et al., 2003] se empleó un modelo de regresión por mínimos cuadrados parciales capaz de predecir algunos aromas de los vinos a partir de su composición química. Se determinaron las características sensoriales de 57 vinos rojos españoles añejados. La composición química de los vinos fue determinada por GC-MS, analizándose un total de 69 compuestos. Ambas matrices, sensorial y química, fueron simplificadas agrupando y reorganizando términos sensoriales correlacionados, excluyendo aromas secundarios o débiles y obteniéndose una matriz final de 18 términos sensoriales y 27 compuestos químicos. Se encontraron modelos satisfactorios para 9 atributos sensoriales que explicaron más del 45% de la varianza original. Dados los altos valores de coeficiente de correlación entre las variables medidas y las predichas, se confirma la existencia de relaciones multivariadas complejas entre aromas y compuestos. En general, los descriptores agradables fueron positivamente correlacionados con compuestos químicos de aromas agradables y negativamente correlacionados con los que mostraban aromas menos favorables.

En [Aishima, 2004] se presenta un estudio que consiste en comparar los perfiles de GC-MS y las respuestas de una nariz electrónica, con los datos sensoriales descriptivos del aroma de salsa de soya. Se encontró que 15 atributos sensoriales descriptivos fueron bien correlacionados con los perfiles GC-MS por análisis de regresión PLS. Se mostró la alta predicción de los modelos PLS, usando 20 picos altamente contribuyentes para predecir cada atributo. Por otro lado la predicción de las respuestas de la nariz electrónica para los atributos sensoriales fue insatisfactoria excepto para dos atributos.

En [Schulbach et al., 2004] se busca realizar comparaciones entre la respuesta sensorial humana y los datos de GC; sin embargo, se argumenta que asociar el análisis sensorial con los datos de GC es una tarea difícil, porque el impacto sensorial de un compuesto generalmente no está asociado al tamaño de su pico en el cromatograma. Las muestras empleadas en este estudio fueron recolectadas de 4 cultivos comerciales de fresas. Éstas fueron analizadas por un grupo de 16 panelistas expertos, donde los atributos a evaluar en una escala de 0 a 100 fueron los siguientes: verde, frutal, floral, caramelo, pera, dulce, ácido y frutilla.

En el estudio realizado se encontró que solo el 20% de los picos detectados tenían actividad aromática y el 50% de los compuestos aromático-activos no fueron detectados. Para resolver este problema, se emplean métodos estadísticos como PLS para seleccionar picos asociados con la respuesta sensorial. Adicionalmente, el impacto sensorial de algunos compuestos puede ser enmascarado por otros compuestos, resaltado por otros o, la combinación de compuestos puede producir una respuesta sensorial enteramente diferente del carácter sensorial de los constituyentes. Algunos atributos sensoriales fueron exitosamente descritos por el análisis instrumental tales como verde, cítrico y dulce. Mientras que otros, frutal y floral, no fueron bien correlacionadas con compuestos que tuvieran dichos aromas.

Las razones principales para correlacionar datos sensoriales e instrumentales son: predecir la preferencia del consumidor, optimizar las condiciones de prueba para el análisis instrumental u obtener un mejor entendimiento de las diferencias de aromas y olores en los productos alimenticios [Schulbach et al., 2004]. La construcción de una ecuación polinomial de atributos sensoriales (variables dependientes) y datos instrumentales (variables independientes) ayuda a predecir o estimar perfiles sensoriales a partir de un conjunto de medidas instrumentales objetivas [Adhikari et al., 2003]. No obstante, hasta el presente ningún método puede ilustrar de forma precisa la composición del aroma de una muestra. Sin embargo, se espera que mediante la integración de variables sensoriales e instrumentales, eventualmente se logre evitar el gasto de tiempo y dinero en perfiles sensoriales. [François et al., 2008, Lindinger et al., 2008]. Además de las razones interpretativas expuestas arriba, otro fin perseguido es la automatización, es decir, usar medidas instrumentales como alternativa a las evaluaciones sensoriales, para concluir que las medidas instrumentales pueden ser usadas como una herramienta para establecer las características aromáticas de un producto en el marco de un programa de control de calidad [Lassoued et al., 2008].

Con respecto al análisis de aromas de café, en [Ribeiro et al., 2009] se reporta la investigación de la relación entre los compuestos volátiles provenientes de café tostado y ciertos atributos sensoriales, incluyendo cuerpo, sabor, limpieza en taza y calidad. Los análisis sensoriales se realizaron sobre 58 muestras de café de Brasil, acompañados de medidas de GC-MS. Los perfiles cromatográficos obtenidos fueron alineados mediante el algoritmo COW, posteriormente la selección de los picos para cada modelo de regresión fue realizada aplicando un algoritmo de selección de predictores significativos. Una vez alineados, los perfiles cromatográficos fueron usados para desarrollar modelos para cada atributo sensorial usando PLSR. Los resultados mostraron la factibilidad de aplicar esta metodología en tiempo real o en aplicaciones de rutina, que requieran predecir la calidad sensorial de muestras de café Arábigo Brasileño.



## 4. Marco experimental

En este capítulo se presentan las bases de datos empleadas y la metodología para el análisis de datos que se resume en el diagrama de bloques de la figura 4.1.

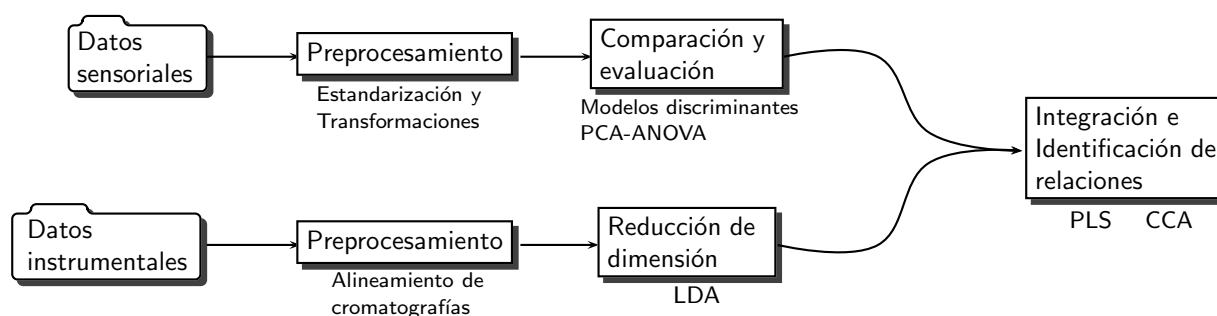


FIGURA 4.1. Diagrama de bloques de la metodología de análisis de datos

### 4.1. Bases de datos

Debido al interés comercial e industrial de los productos alimenticios y al alto costo de las evaluaciones sensoriales y las pruebas instrumentales, resulta muy difícil contar con conjuntos de datos públicos que incluyan ambos tipos de mediciones y que cumplan con un diseño de experimentos determinado. Para este trabajo, la base de datos principal fue suministrada por una fábrica de café de la región. Adicionalmente, se usaron dos conjuntos de datos públicos que incluyen únicamente evaluaciones sensoriales. Estos dos últimos conjuntos corresponden a evaluaciones de productos de papa y de queso. A continuación se presenta una descripción concisa de las bases de datos.

### 4.2. Protocolo de adquisición de datos sensoriales

La obtención de datos sensoriales debe realizarse siguiendo un protocolo de adquisición de las muestras basado en las siguientes Normas Técnicas Colombianas. La NTC 4129 [ICONTEC, 1997b] brinda una descripción de los métodos de selección, entrenamiento y seguimiento de los evaluadores. Para realizar la selección de los catadores se deben explorar los siguientes aspectos: interés y motivación, actitudes hacia los alimentos, conocimiento de los aspectos del producto y aptitud de concentración, buena salud, buena capacidad de comunicación, disponibilidad de tiempo, interés y motivación, etc. Las pruebas de selección a efectuar consisten en la determinación de deterioro en las capacidades sensoriales y agudeza visual y evaluación potencial del candidato para describir y comunicar percepciones sensoriales. Se procede a realizar las pruebas de correspondencia, en el cual se presenta una serie de muestras de cada sabor básico para la etapa de familiarización. Luego estas mismas

muestras son presentadas con diferente codificación y se pide a cada uno de ellos que comparen con el grupo original y describan la sensación experimentada. En [ICONTEC, 1997b] se presentan ejemplos de materiales y sus concentraciones para las pruebas de correspondencia. Otro tipo de prueba que debe efectuarse es el de agudeza y capacidad de discriminación del gusto, olor, textura y color, el cual se lleva a cabo analizando diferentes productos. Una vez efectuada la prueba anterior, se realiza la prueba de capacidad descriptiva, en la cual se determina la capacidad de los candidatos para describir percepciones sensoriales a través de 2 exámenes: estímulos odoríferos y de texturas. En el primero se evalúan estímulos 5 a 10 olfatorios relacionados con el producto a evaluar, en el segundo, se suministra una serie de productos al azar para describir sus características. Otros detalles sobre estas pruebas se encuentran en [ICONTEC, 1997b]

El entrenamiento debe efectuarse en un medio adecuado, tal como lo indica la NTC 3925 [ICONTEC, 1996]. También es útil entrenar a los evaluadores en conocimiento básico de los productos a evaluar. Esto se hace brindándoles información acerca del proceso de fabricación y organizando visitas a la planta de producción. En el entrenamiento para la detección y reconocimiento de sabores y olores, inicialmente debe hacerse las pruebas de correspondencia, reconocimientos, comparación por pares, test triangular y duo-trio para un correcto entrenamiento tanto a bajas concentraciones como a altas. Los diferentes métodos de entrenamiento del panel sensorial se describen en las normas NTC 2679 [ICONTEC, 1997a], sobre el ensayo "A" y "NO A"; NTC 2680 [ICONTEC, 1990], sobre el ensayo de comparación por pares y la NTC 2681 [ICONTEC, 2006], sobre la prueba triangular. En cuanto al uso de escalas, los catadores deben conocer previamente conceptos de escalas de valoración, clasificación, intervalos y proporciones con el fin de atribuir magnitudes significativas a las muestras. Respecto al uso de los descriptores, debe establecerse un vocabulario para describir las características sensoriales en términos particulares que permitan diferenciar las muestras, el cual debe contener como mínimo 10 descriptores. Los atributos seleccionados para la realización de las evaluaciones deben estar contemplados en la norma NTC 2758 [ICONTEC, 2002].

La metodología de evaluación de muestras que se recomienda cuando la aplicación involucra la correlación de los datos sensoriales con información instrumental, es el análisis cuantitativo descriptivo, descrito en la NTC 4883 [ICONTEC, 2000]. Para este análisis, los catadores que no puedan reconocer o describir adecuadamente y en forma correcta el 70% de las muestras de control, se deben considerar incompetentes para este tipo de prueba. Los catadores deben evaluar 6 muestras usando el vocabulario y la hoja de puntaje elaborada previamente. En cuanto al número de catadores se requieren mínimo 3 entrenados y expertos. Para validar resultados con un número inferior se debe hacer evaluación por triplicado de cada muestra. La escala de acuerdo al descriptor indica la intensidad del atributo cuya calificación aumentará o disminuirá de acuerdo a su presencia en el producto. El reporte del ensayo debe contener: objetivo del análisis, tipo de café analizado, método de muestreo, método de preparación de la muestra, condiciones de tostación y reporte de resultados en una tabla o gráfico.



#### 4.2.1. Base de datos sensoriales e instrumentales de café

Este conjunto está compuesto por dos tipos de datos medidos en muestras de café tostado: evaluaciones sensoriales y medidas químicas obtenidas de la cromatografía de gases de cada producto.

Los datos sensoriales comprenden las medidas de 9 atributos evaluados por un panel de expertos compuesto por tres catadores (C1, C2 y C3). Las muestras de café fueron tostadas a tres temperaturas diferentes que denominaremos T1, T2 y T3. Adicionalmente, para cada condición de tostación se evaluaron muestras de 4 regiones de Colombia con el objetivo de caracterizar cada región con respecto a los aromas desarrollados en el proceso de tostación. En total se obtuvieron 6 réplicas de cada combinación de región y temperatura por cada catador. Los 9 atributos se listan en la tabla 4.1.

TABLA 4.1. Atributos sensoriales usados en el análisis descriptivo de las muestras de café tostado

Atributos Olfativos		
Floral	Dulce	Herbal
Frutal	Cereal	Cigarrillo
Tostado	Madera	Cítrico

Adicionalmente, los datos de cromatografía de gases comprenden señales de 10277 puntos, todas de igual longitud, para la misma cantidad de muestras y réplicas que la base de datos sensorial. Estas señales fueron obtenidas en dos periodos diferentes, una primera parte de las señales fue obtenida en un periodo de 3 semanas en el mes de junio de 2009 y una segunda parte fue obtenida en un periodo de 2 semanas del mes de enero de 2010. Debido a diferentes condiciones de configuración del instrumento de cromatografía, las señales obtenidas en 2009 y 2010 presentan diferencias importantes que deben ser resueltas antes de continuar con el análisis (ver sección 3.2). El cromatógrafo proporciona además información de cada pico de la señal tal como: altura, área, tiempo de retención, compuesto asociado al pico, probabilidad de que ese compuesto sea el detectado por el equipo, #CAS y área de cada pico en terminos porcentuales respecto al área total bajo la curva.

#### 4.2.2. Base de datos sensoriales de papa

Esta base de datos <sup>1</sup> está compuesta por 6 variedades de papa de Dinamarca [Thybo and Martens, 1998]. El panel sensorial está compuesto por 7 mujeres y 3 hombres (C1–C10) de edades entre los 20 y 45 años, compuesto por técnicos y personal de apoyo del instituto *Danish Institute of Agricultural Sciences*. El panel sensorial fue entrenado durante 6 horas previas a la evaluación sensorial. Los catadores habían tenido experiencia en evaluar la textura de papas cocinadas en evaluaciones previas, con 40 horas aproximadamente de experiencia. Los atributos sensoriales de textura (Tabla 4.2) fueron evaluados en una escala semiestructurada de 1-9 puntos. Cada catador evaluó los seis productos en 10 réplicas.

<sup>1</sup>Disponible en: <http://www.models.life.ku.dk/research/data/>

TABLA 4.2. Atributos sensoriales usados en el análisis descriptivo de la muestras de papa

Atributos gustativos	Atributos visuales
Dureza	Reflejo de la superficie
Firmeza	
Adherencia	
Elasticidad	
Granuloso	
Harinoso	
Húmedo	
Masticabilidad	

#### 4.2.3. Base de datos sensoriales de queso

Esta base de datos <sup>1</sup> comprende la evaluación sensorial de 23 atributos sensoriales para nueve tipos de queso crema diferentes [Bro et al., 2008]. Estas evaluaciones sensoriales fueron realizadas en tres repeticiones. El panel sensorial estaba compuesto por ocho cata-dores, que fueron entrenados durante 5 sesiones con los productos empleados, y los atributos sensoriales fueron escogidos por sugerencia del líder del panel con base en un consenso entre los panelistas. Los 32 atributos se listan en la tabla 4.3. Para la evaluación de los atributos se usó una línea horizontal de escala de 15cm no estructurada. En cada sesión se evaluaron 5 de los 9 productos.

TABLA 4.3. Atributos sensoriales usados en el análisis descriptivo de la muestras de queso crema

Atributos Olfativos	Atributos visuales	Atributos gustativos
Aroma a crema (O-crema)	Color blanco (V-blanco)	Firme (G-Firme)
Aroma ácido (O-Acido)	Color gris (V-gris)	Fundido (G-Fundi)
Aroma a mantequilla (O-Manteq)	Color amarillo (V-Amarill)	Resistencia (G-Resist)
Aroma a leche vieja (O-Leche vieja)	Color verde (V-Verde)	granuloso (G-Granul)
	Granulos (V-Granul)	Calcáreo (G-Calcáreo)
	Brillante (V-Brillan)	Sabor a crema (G-Crema)
		Sabor graso (G-Graso)
Atributo táctil		Sabor a Mantequilla (G-Manteq)
Resistencia (T-Resist)		Salado (G-Sal)
		Vinagre (G-Vinag)
		Dulce (G-Dulce)

#### 4.3. Análisis exploratorio de la uniformidad de los evaluadores

Idealmente, la medición de una misma variable usando diferentes instrumentos debería producir idénticas evaluaciones o por lo menos con variaciones pequeñas. De lo contrario, se presume que uno o más de los instrumentos está(n) descalibrado(s). Análogamente, en el caso de las evaluaciones sensoriales las variaciones o discrepancias en las medidas de un mismo producto son indeseables. Además de la uniformidad, las evaluaciones deberían ser

confiables, es decir, que las medidas reflejen fielmente la presencia y el grado de los atributos en la muestra. Con el fin de asegurar la uniformidad y la confiabilidad mencionadas, se propone en la literatura la aplicación de una serie de transformaciones y correcciones. A continuación se presenta la aplicación de las correcciones presentadas en las secciones 3.1.3–3.1.8 sobre los datos descritos en 4.1.

#### 4.3.1. Interpretación de datos sensoriales usando análisis de varianza

Un modelo ANOVA de una vía balanceado se emplea para evaluar si los catadores actúan como un factor que influye significativamente en la respuesta de los atributos medidos. Los resultados de las pruebas de hipótesis se encuentran en el apéndice A. Las tablas de éste apéndice corresponden a los resultados de las pruebas de hipótesis de ANOVA para cada atributo. Al inicio de cada tabla se observan los resultados del modelo ANOVA de una vía para los catadores. De todos los atributos, el único que no muestra influencia del catador es el atributo tostado, pues se observa que su valor-p es 0.31, lo que indica que el factor catador no es significativamente diferente para este atributo. Para los otros 8 atributos, el efecto del catador sí es significativo a un nivel de 0.05. Adicionalmente, se realizaron pruebas de comparación múltiple usando el criterio de diferencia de Tukey. Los resultados de la prueba, mostrados en la tabla 4.4, indican que, para los atributos floral, frutal, dulce, cereal y cítrico, los tres catadores son significativamente diferentes a niveles de 0.05 y 0.01. Los resultados de esta prueba muestran la necesidad de emplear las transformaciones y correcciones, tal que se elimine el desacuerdo entre los catadores. Efectivamente, para los atributos frutal, cereal y cítrico mostrados en la tabla 4.4, las diferencias entre los pares de catadores son significativas a un nivel de 0.01. Se confirma además que, como ocurrió para la prueba de ANOVA de una vía, las evaluaciones del atributo tostado de todos los catadores no tienen diferencias significativas.

TABLA 4.4. Estadística de Tukey para la prueba de diferencia entre los catadores de los atributos

Catadores	Floral	Frutal	Tostado	Dulce	Cereal	Madera	Herbal	Cigarr.	Cítrico
1 - 2	-1.52**	-1.96**	0.22	-1.76**	-1.65**	-1.29**	-0.70**	-1.17**	-1.89**
1 - 3	0.23*	-0.33**	0.16	-1.52**	-1.04**	-0.12	-0.37*	0.09	-0.47**
2 - 3	1.75**	1.63**	-0.06	0.24*	0.61**	1.17**	0.33	1.26**	1.41**

\* diferentes a un nivel de significancia de 0.05

\*\* diferentes a un nivel de significancia de 0.01

Se ha mencionado con anterioridad que ANOVA permite evaluar el efecto de los diferentes factores experimentales, que en este análisis comprenden temperaturas, regiones y catadores. Como se puede ver en las tablas A.1-A.9 de ANOVA que se encuentran en el apéndice A, según el valor F el efecto del catador para los datos sin corregir es casi siempre el efecto dominante en todos los atributos, seguido por el efecto de la temperatura y la interacción temperatura×catador. Se puede observar que estos son los tres factores dominantes para todos los atributos sobre los datos sin corrección.

Las correcciones funcionan muy bien para eliminar el efecto del catador. La corrección usando análisis de Procrustes incide en los efectos y se observa que los tres efectos dominantes son en general temperatura, región y una interacción temperatura×región o temperatura×catador. Para la corrección de Brockhoff, los efectos dominantes son la temperatura y la región.

Para los datos sin corrección, el efecto del catador es obviamente un efecto significativamente diferente, mientras que los dos métodos de corrección empleados logran que este efecto no sea significativamente diferente a un nivel de 0.05, e incluso para algunos atributos, la interacción temperatura×catador tampoco es significativamente diferente, de tal modo que se logran resolver las diferencias de nivel producidas por los catadores, y se reduce el efecto de las diferencias de escala y desacuerdo que están representadas por la interacción temperatura×catador. A continuación se relacionan los atributos que lograron ser corregidos en dicho aspecto. La transformación de Procrustes eliminó esta interacción para los atributos, floral, frutal, dulce, madera, herbal y cigarrillo. El modelo de evaluador con su corrección de escalado multiplicativo, elimina la interacción temperatura×catador para los atributos floral, tostado, dulce, cereal, madera y cigarrillo, mientras la modalidad de escalado aditivo elimina dicha interacción sobre los atributos frutal, tostado, dulce, cereal, cigarrillo y cítrico.

#### **4.3.2. Interpretación de datos sensoriales usando PCA-ANOVA**

El objetivo de usar esta metodología propuesta por [Luciano and Næs, 2009] es poder interpretar la concordancia del panel de catación para cada atributo sensorial y consiste en combinar PCA con ANOVA para darle un enfoque multivariado que proporciona información sobre el efecto de los catadores y los productos, además de su interacción, en la estructura multivariada. La ventaja de usar PCA-ANOVA es su simplicidad y que proporciona pruebas de significancia basadas en ANOVA y herramientas visuales para mejorar la interpretación de las pruebas de hipótesis (efectos principales e interacciones). Adicionalmente, esta metodología puede ser extendida a situaciones con varios factores de diseño y estructuras complejas de las réplicas.

Inicialmente se realiza una inspección visual de los datos usando la transformación obtenida del análisis de componentes principales. Se visualizaron todas las posibles combinaciones de parejas coordenadas de las cuatro primeras componentes principales. Debido a que no se observaron diferencias notables entre sus distribuciones, se muestran aquí únicamente las gráficas de dispersión de componente 1 contra componente 2 y componente 3 contra componente 4.

En las figuras 4.2 - 4.4 se ilustran las gráficas de dispersión de los datos sobre parejas de las cuatro primeras componentes. En estas figuras se pueden identificar las tres condiciones o factores experimentales estudiados; a saber, tres temperaturas, cuatro regiones y tres catadores. Cada uno de estos últimos es un instrumento de medición para evaluar los 9 atributos (ver sección 4.2.1), bajo las diferentes condiciones de temperatura y región. En la figura 4.2 se muestran las cuatro primeras componentes principales (en adelante PC) de los

datos sin corregir medidos por los catadores. Se usan los colores de los puntos para hacer referencia a las regiones y la forma de los mismos para indicar las temperaturas. En esta figura se puede observar que las temperaturas forman agrupaciones sobre las dos primeras componentes principales; es decir, aparentemente la dispersión entre temperaturas es mayor que la dispersión intra temperaturas. Por otro lado, se observa que las cuatro regiones presentan una dispersión intra clase alta, mientras que la dispersión entre dichas clases es pequeña, pues se observa que las distribuciones de las clases están casi concéntricas sobre las cuatro primeras componentes.

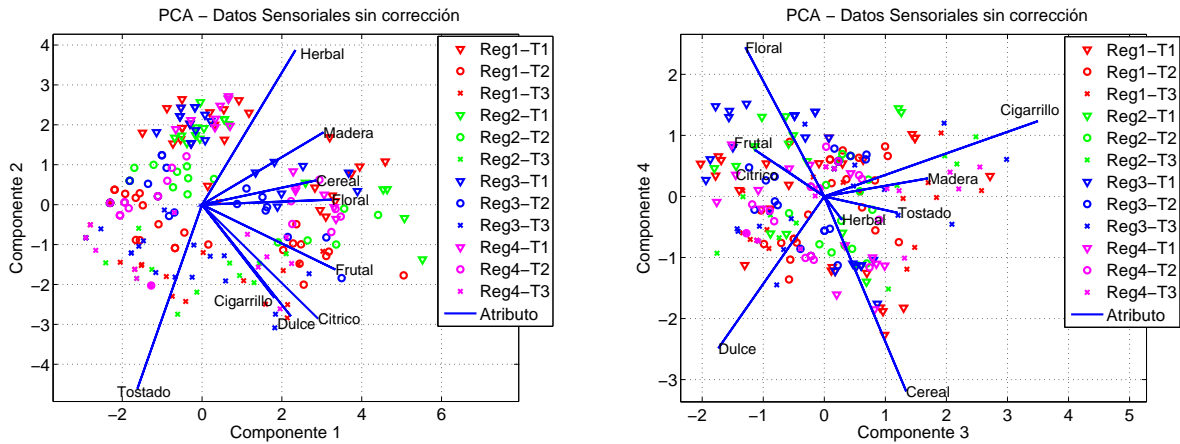


FIGURA 4.2. Biplot PCA para los datos sensoriales originales (muestras etiquetadas por regiones y temperaturas)

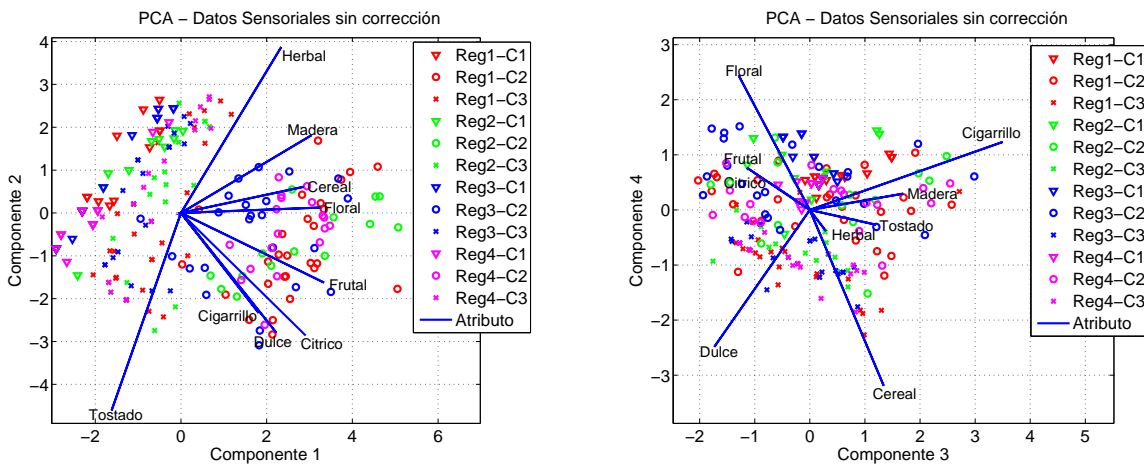


FIGURA 4.3. Biplot PCA para los datos sensoriales originales (muestras etiquetadas por regiones y catadores)

La gráfica de dispersión en la figura 4.3 muestra la misma representación PCA de los datos sin corregir, pero en este caso, los colores representan las regiones y el tipo de los puntos

representa los tres catadores. Es evidente que las dos primeras componentes capturan la varianza debida a las diferencias individuales entre los catadores. Se observan dos agrupamientos en la gráfica de las dos primeras PC, uno formado por el catador 2 y el otro por los catadores 1 y 3. Esta gráfica sugiere que los atributos floral, cereal, frutal, cítrico dulce y cigarrillo reflejan estas diferencias entre el catador 2 y los demás. De la misma forma que la figura anterior, no se observan grupos correspondientes a las regiones; de hecho, los colores están muy mezclados.

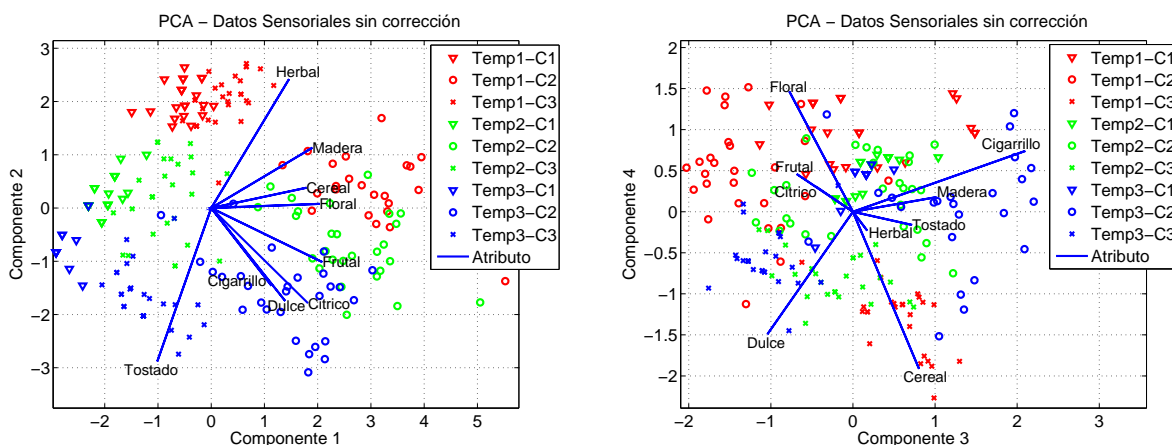


FIGURA 4.4. Biplot PCA para los datos sensoriales originales (muestras etiquetadas por temperaturas y catadores)

En la figura 4.4 se cambian nuevamente los colores y tipo de puntos, donde estos últimos representan los catadores y los colores representan las temperaturas. En la figura de las dos primeras componentes se observan más claramente las agrupaciones formadas por las temperaturas, y también las sub-agrupaciones formadas por la diferencia entre el catador 2 con los otros dos catadores. En la gráfica de las componentes 3 y 4 se siguen formando agrupaciones entre temperaturas aunque no tan definidas; además, se siguen evidenciando las diferencias entre los catadores en la evaluación de las temperaturas. También es de resaltar que el catador con mayor dispersión es el catador 2.

Las gráficas de dispersión (figuras 4.2 - 4.4) de los datos originales transformados con PCA, sugieren que la incertidumbre del instrumento, es decir, los catadores, es mayor que la incertidumbre del fenómeno, en este caso las diferencias entre regiones y las diferencias entre temperaturas.

El análisis de las relaciones encontradas entre atributos empleando PCA, puede complementarse con la matriz de correlación entre dichas variables (ver Tabla 4.5). Esta matriz muestra una alta dependencia entre los atributos floral y frutal, con una correlación positiva de 0.79; lo cual indica también que estos atributos son directamente proporcionales. Otra alta correlación, de 0.76, se presenta entre los atributos herbal y tostado, pero de signo negativo, lo cual implica proporcionalidad inversa. Adicionalmente, el atributo frutal

también presenta una gran correlación con el atributo cítrico, con un valor positivo de 0.76. Se puede destacar que los índices de correlación mencionados tienen una clara interpretación de dependencia. Asimismo, se observa en la matriz, con negrita, la alta correlación entre otros pares de atributos.

TABLA 4.5. Matriz de correlación de atributos sensoriales

	Floral	Frutal	Tostado	Dulce	Cereal	Madera	Herbal	Cigarrillo	Cítrico
Floral	1.0000	<b>0.7919</b>	-0.3881	0.3778	0.4124	0.5624	0.4604	0.2647	<b>0.6154</b>
Frutal	<b>0.7919</b>	1.0000	-0.1447	0.5830	0.5256	0.5134	0.2594	0.3433	<b>0.7634</b>
Tostado	-0.3881	-0.1447	1.0000	0.0340	-0.2543	-0.4882	<b>-0.7610</b>	0.2123	0.0437
Dulce	0.3778	0.5830	0.0340	1.0000	0.3472	0.1590	0.0660	0.2683	<b>0.6324</b>
Cereal	0.4124	0.5256	-0.2543	0.3472	1.0000	<b>0.6405</b>	0.4910	0.2974	0.4095
Madera	0.5624	0.5134	-0.4882	0.1590	<b>0.6405</b>	1.0000	0.5916	0.3785	0.3723
Herbal	0.4604	0.2594	<b>-0.7610</b>	0.0660	0.4910	0.5916	1.0000	0.0929	0.1122
Cigarrillo	0.2647	0.3433	0.2123	0.2683	0.2974	0.3785	0.0929	1.0000	0.4413
Cítrico	<b>0.6154</b>	<b>0.7634</b>	0.0437	<b>0.6324</b>	0.4095	0.3723	0.1122	0.4413	1.0000

#### 4.3.3. Evaluación de las correcciones del efecto de catador

En esta sección se presentan los resultados de las pruebas de significancia ANOVA, aplicadas sobre las cuatro primeras componentes principales que fueron escogidas por resumir al menos el 80% de la varianza presente en los datos, antes y después de las correcciones.

En la tabla 4.6 se presentan los resultados de pruebas de hipótesis ANOVA, tal como se muestra en la tabla 3.3, para los datos antes y después de las diferentes correcciones o transformaciones, a saber, transformación de Procrustes y correcciones usando el modelo de evaluador para sus dos modalidades, escalado multiplicativo y aditivo. A la derecha del nombre de la corrección aplicada se presenta la varianza explicada de los datos. Particularmente, para los datos originales, la primera componente explica el 46.46% de la varianza. Se observa que los efectos principales de temperatura, catador, región y las interacciones temperatura×catador y catador×región, son significativamente diferentes a un nivel de 0.01 mientras que las interacciones temperatura×región y temperatura×catador×región no son significativamente diferentes. Sin embargo, según el valor  $F$ , los efectos dominantes son la temperatura y el catador.

Para la transformación usando el análisis de Procrustes (tabla 4.6) se observa que el efecto principal del catador deja de ser un efecto significativamente diferente, lo cual indica que se eliminan las diferencias de nivel mencionadas en la sección 3.1.6; mientras que los efectos principales de temperatura y región, y los efectos de interacción siguen siendo significativamente diferentes, incluso el efecto temperatura×región y temperatura×catador×región. El efecto dominante en los datos después de aplicar esta transformación es el efecto de la temperatura.

TABLA 4.6. Comparación de corrección usando PCA-ANOVA (componente 1)

Factor	Suma de cuadrados	grados de libertad	Error cuadrático medio	F	valor p
<b>Sin corrección (<math>\sigma^2</math> explicada 46.46%)</b>					
temp	153.44	2	76.72	228.05	0.0000
catador	813.26	2	406.63	1208.72	0.0000
región	11.40	3	3.80	11.3	0.0000
temp*catador	10.45	4	2.61	7.77	0.0000
temp*región	2.38	6	0.40	1.18	0.3178
catador*región	7.50	6	1.25	3.72	0.0015
catador*temp*región	5.32	12	0.44	1.32	0.2088
Error	78.72	234	0.34		
Total	1124.79	269			
<b>Procrustes (<math>\sigma^2</math> explicada 45.96%)</b>					
temp	887.20	2	443.60	868.39	0.0000
catador	0.19	2	0.09	0.19	0.8305
región	5.49	3	1.83	3.58	0.0146
temp*catador	9.53	4	2.38	4.66	0.0012
temp*región	12.23	6	2.04	3.99	0.0008
catador*región	8.82	6	1.47	2.88	0.0100
temp*catador*región	19.35	12	1.612	3.16	0.0003
Error	119.53	234	0.511		
Total	1112.80	269			
<b>Modelo de evaluador - escalado multiplicativo (<math>\sigma^2</math> explicada 31.02%)</b>					
temp	492.33	2	246.16	334.03	0.0000
catador	0.06	2	0.03	0.04	0.9613
región	14.05	3	4.68	6.60	0.0003
temp*catador	17.15	4	4.29	6.04	0.0001
temp*región	7.83	6	1.31	1.84	0.0923
catador*región	15.37	6	2.56	3.61	0.0019
temp*catador*región	15.24	12	1.27	1.79	0.0507
Error	181.29	246	0.71		
Total	750.93	269			
<b>Modelo de evaluador - escalado aditivo (<math>\sigma^2</math> explicada 51.18%)</b>					
temp	157.38	2	78.69	231.68	0.0000
catador	820.23	2	410.11	1207.48	0.0000
región	11.56	3	3.85	11.35	0.0000
temp*catador	10.48	4	2.62	7.71	0.0000
temp*región	2.40	6	0.40	1.18	0.3202
catador*región	1.04	6	0.17	0.51	0.8008
temp*catador*región	5.38	12	0.45	1.32	0.2082
Error	79.48	234	0.34		
Total	1128.95	269			



La corrección de escala de Brockhoff, para su versión de escalado multiplicativo (tabla 4.6), elimina el efecto del catador mientras conserva los demás efectos diferentes a un nivel de significancia de 0.01, exceptuando los efectos de interacción temperatura×región, que antes de corregir no era significativamente diferente, y temperatura×catador×región que sí lo era.

En contraste, la corrección de los datos usando el escalado aditivo de Brockhoff, elimina todos los efectos de interacción que representan las diferencias de escala y desacuerdo entre los catadores, como se muestra en la sección 3.1.6. No obstante, se observa que los efectos dominantes siguen siendo los efecto principales de catador y de temperatura, igual que en los datos sin corregir.

Para el segundo componente PCA (ver tabla 4.7), se observa que para los datos sin corregir los efectos principales de temperatura, catador y región, son significativamente diferentes, de la misma forma que todos los efectos de interacción. Sin embargo, los efectos dominantes siguen siendo la temperatura y el catador. Con la corrección Procrustes se elimina el efecto del catador mientras se conservan las diferencias significativas para los otros efectos principales e interacciones temperatura×región y catador×región. Por otro lado, la corrección con el modelo del evaluador con escalado multiplicativo, elimina el efecto principal del catador y de interacciones de temperatura×región y catador×región. La corrección usando el escalado aditivo eliminó el efecto de la interacción catador×región, la cual representa las diferencias de escala y desacuerdo de los catadores al medir las regiones, mientras los demás efectos presentes en los datos sin corrección permanecieron significativamente diferentes.

Para el tercer PC la varianza presente es muy poca, alrededor del 10%. Sin embargo, para los datos sin corregir, todos los efectos principales y de interacción son significativamente diferentes. No obstante, según el valor  $F$ , el efecto dominante es la interacción temperatura×catador. Los resultados de ANOVA para la transformación de Procrustes indican que el único efecto significativamente diferente es la interacción temperatura×región y que la mayor parte de la varianza está presente en el error. El escalado multiplicativo del modelo de evaluador presenta resultados similares, donde los efectos significativamente diferentes son el de región, y las interacciones temperatura×región y catador×región. El modelo del evaluador usando escalado aditivo muestra como efecto dominante a la interacción temperatura×catador, seguido por los efecto principales de temperatura y catador; además, el efecto de región y el efecto de interacción temperatura×región, también son significativamente diferentes. Un análisis similar puede hacerse para los resultados de ANOVA para la cuarta componente PCA.

TABLA 4.7. Comparación de corrección usando PCA-ANOVA (componente 2)

Factor	Suma de cuadrados	grados de libertad	Error cuadrático medio	F	valor p
<b>Sin corrección (<math>\sigma^2</math> explicada 68.46%)</b>					
temp	347.10	2	173.55	675.98	0.0000
catador	54.35	2	27.17	105.84	0.0000
región	3.42	3	1.14	4.44	0.0047
temp*catador	24.50	4	6.12	23.86	0.0000
temp*región	6.02	6	1.00	3.91	0.0010
catador*región	4.81	6	0.80	3.12	0.0058
temp*catador*región	7.55	12	0.63	2.45	0.0050
Error	60.08	234	0.26		
Total	532.53	269			
<b>Procrustes (<math>\sigma^2</math> explicada 63.25%)</b>					
temp	61.64	2	30.82	32.92	0.0000
catador	0.13	2	0.06	0.07	0.9334
región	53.27	3	17.76	18.97	0.0000
temp*catador	34.81	4	8.70	9.30	0.0000
temp*región	9.33	6	1.55	1.66	0.1317
catador*región	29.39	6	4.90	5.23	0.0000
temp*catador*región	11.05	12	0.92	0.98	0.4652
Error	219.08	234	0.94		
Total	418.40	269			
<b>Modelo de evaluador - escalado multiplicativo (<math>\sigma^2</math> explicada 45.02%)</b>					
temp	53.65	2	26.82	33.28	0.0000
catador	0.07	2	0.03	0.04	0.9597
región	21.19	3	7.06	8.76	0.0000
temp*catador	5.05	4	1.26	1.57	0.1842
temp*región	4.82	6	0.80	1.00	0.4286
catador*región	54.36	6	9.06	11.24	0.0000
temp*catador*región	10.43	12	0.87	1.08	0.3791
Error	188.61	234	0.81		
Total	338.98	269			
<b>Modelo de evaluador - escalado aditivo (<math>\sigma^2</math> explicada 74.65%)</b>					
temp	346.61	2	173.30	659.85	0.0000
catador	58.75	2	29.38	111.85	0.0000
región	3.53	3	1.18	4.48	0.0044
temp*catador	24.71	4	6.18	23.52	0.0000
temp*región	6.06	6	1.01	3.85	0.0011
catador*región	1.17	6	0.19	0.74	0.6182
temp*catador*región	7.62	12	0.63	2.42	0.0057
Error	61.46	234	0.26		
Total	531.32	269			

TABLA 4.8. Comparación de corrección usando PCA-ANOVA (componente 3)

Factor	Suma de cuadrados	grados de libertad	Error cuadrático medio	F	valor p
<b>Sin corrección (<math>\sigma^2</math> explicada 78.90%)</b>					
temp	7.32	2	3.66	13.16	0.0000
catador	6.44	2	3.22	11.58	0.0000
región	3.26	3	1.09	3.91	0.0094
temp*catador	144.20	4	36.05	129.74	0.0000
temp*región	4.37	6	0.73	2.62	0.0176
catador*región	6.50	6	1.08	3.90	0.0010
temp*catador*región	6.41	12	0.53	1.92	0.0325
Error	65.02	234	0.28		
Total	252.93	269			
<b>Procrustes (<math>\sigma^2</math> explicada 74.81%)</b>					
temp	1.44	2	0.72	0.76	0.4671
catador	0.09	2	0.05	0.05	0.9518
región	7.18	3	2.39	2.54	0.0569
temp*catador	2.68	4	0.67	0.71	0.5851
temp*región	26.84	6	4.47	4.75	0.0001
catador*región	5.16	6	0.86	0.91	0.4856
temp*catador*región	17.39	12	1.45	1.54	0.1108
Error	220.17	234	0.94		
Total	279.96	269			
<b>Modelo de evaluador - escalado multiplicativo (<math>\sigma^2</math> explicada 56.75%)</b>					
temp	0.24	2	0.12	0.12	0.8842
catador	0.32	2	0.16	0.17	0.8456
región	8.46	3	2.82	2.91	0.0352
temp*catador	0.48	4	0.12	0.13	0.9733
temp*región	27.82	6	4.64	4.79	0.0001
catador*región	16.05	6	2.68	2.76	0.0130
temp*catador*región	4.19	12	0.35	0.36	0.9756
Error	226.56	234	0.97		
Total	284.08	269			
<b>Modelo de evaluador - escalado aditivo (<math>\sigma^2</math> explicada 82.91%)</b>					
temp	6.37	2	3.19	11.27	0.0000
catador	7.06	2	3.53	12.48	0.0000
región	3.23	3	1.08	3.81	0.0108
temp*catador	146.22	4	36.56	129.27	0.0000
temp*región	4.39	6	0.73	2.59	0.0190
catador*región	2.02	6	0.34	1.19	0.3112
temp*catador*región	6.52	12	0.54	1.92	0.0326
Error	66.17	234	0.28		
Total	251.49	269			

TABLA 4.9. Comparación de corrección usando PCA-ANOVA (componente 4)

Factor	Suma de cuadrados	grados de libertad	Error cuadrático medio	F	valor p
<b>Sin corrección (<math>\sigma^2</math> explicada 85.86%)</b>					
temp	0.78	2	0.39	2.20	0.1128
catador	98.32	2	49.16	278.23	0.0000
región	3.27	3	1.09	6.18	0.0005
temp*catador	11.64	4	2.91	16.47	0.0000
temp*región	2.33	6	0.39	2.20	0.0440
catador*región	1.93	6	0.32	1.82	0.0966
temp*catador*región	1.06	12	0.09	0.50	0.9148
Error	41.35	234	0.18		
Total	168.46	269			
<b>Procrustes (<math>\sigma^2</math> explicada 83.78%)</b>					
temp	22.92	2	11.46	17.31	0.0000
catador	0.09	2	0.04	0.07	0.9363
región	4.86	3	1.62	2.45	0.0647
temp*catador	6.53	4	1.63	2.47	0.0458
temp*región	15.23	6	2.54	3.83	0.0011
catador*región	5.03	6	0.84	1.27	0.2743
temp*catador*región	5.45	12	0.45	0.69	0.7645
Error	154.92	234	0.66		
Total	217.12	269			
<b>Modelo de evaluador - escalado multiplicativo (<math>\sigma^2</math> explicada 67.94%)</b>					
temp	4.28	2	2.14	2.41	0.0924
catador	0.21	2	0.10	0.12	0.8907
región	4.68	3	1.56	1.75	0.1566
temp*catador	4.35	4	1.09	1.22	0.3017
temp*región	1.83	6	0.31	0.34	0.9134
catador*región	17.11	6	2.85	3.20	0.0048
temp*catador*región	30.97	12	2.58	2.90	0.0009
Error	208.22	234	0.89		
Total	270.77	269			
<b>Modelo de evaluador - escalado aditivo (<math>\sigma^2</math> explicada 87.53%)</b>					
temp	0.72	2	0.36	2.08	0.1279
catador	104.36	2	52.18	301.03	0.0000
región	3.45	3	1.15	6.63	0.0003
temp*catador	10.45	4	2.61	15.08	0.0000
temp*región	2.31	6	0.38	2.22	0.0423
catador*región	0.39	6	0.06	0.37	0.8967
temp*catador*región	1.04	12	0.09	0.50	0.9124
Error	40.56	234	0.17		
Total	168.91	269			

Como se mencionó en el marco teórico (sección 3.1.5), para los efectos del producto, catador y de interacción que son calculados sobre los componentes principales  $T$  de la ecuación (17), se puede mejorar la interpretación de la variabilidad multivariada en los datos, mediante la visualización de los efectos en las gráficas de los componentes principales. Téngase en cuenta que los datos transformados siguen representando los mismos efectos que los datos originales; obteniéndose así un análisis similar al realizado en la sección 4.3.2.

En la figura 4.5 se observan los datos corregidos a partir del análisis de Procrustes y transformados usando PCA. La primera componente principal representa las diferencias entre las temperaturas, que como se comentó anteriormente, es el efecto dominante en esta componente. Los atributos tostado, cigarrillo, floral, frutal y herbal son los atributos que tienen la ponderación más alta para este componente; además, se nota que los atributos tostado y cigarrillo tienen una correlación negativa con los demás atributos mencionados. Los atributos cereal, madera y cítrico son los atributos que más contribuyen a la componente 2. Para las componentes 3 y 4 es evidente que los atributos cítrico y dulce son los de mayor ponderación.

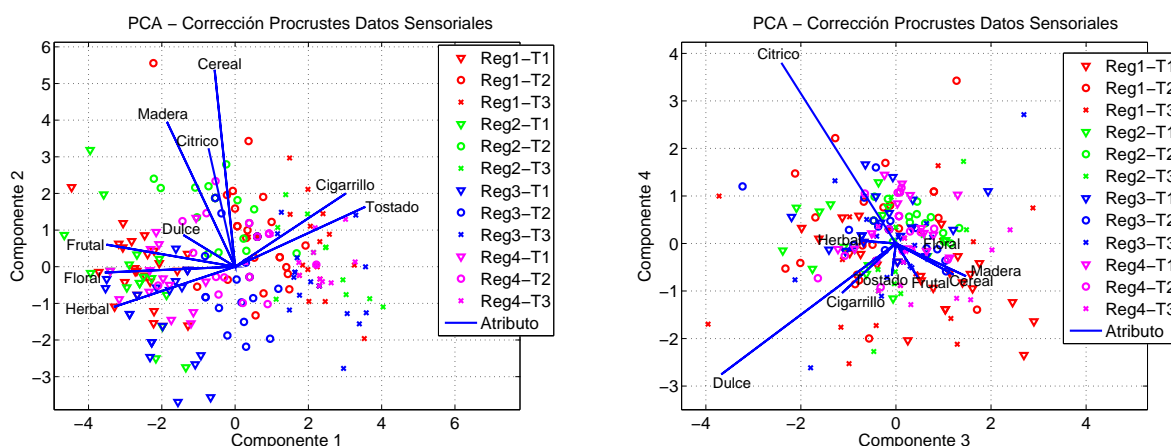


FIGURA 4.5. Biplot PCA. Corrección Procrustes para datos sensoriales de café.

En la figura 4.6 se presentan dos gráficas de dispersión para parejas de las cuatro primeras componentes principales de los datos transformados con la corrección de Brockhoff - escalado multiplicativo. En la gráfica de las dos primeras componentes se observa el traslape entre las regiones, con la particularidad de que la dispersión de la región 4 es menor que para las otras tres regiones. Después de la transformación de los datos se siguen formando grupos por temperaturas. Las componentes 3 y 4 forman una distribución elongada de los datos, donde la región 1 es la que muestra mayor dispersión.

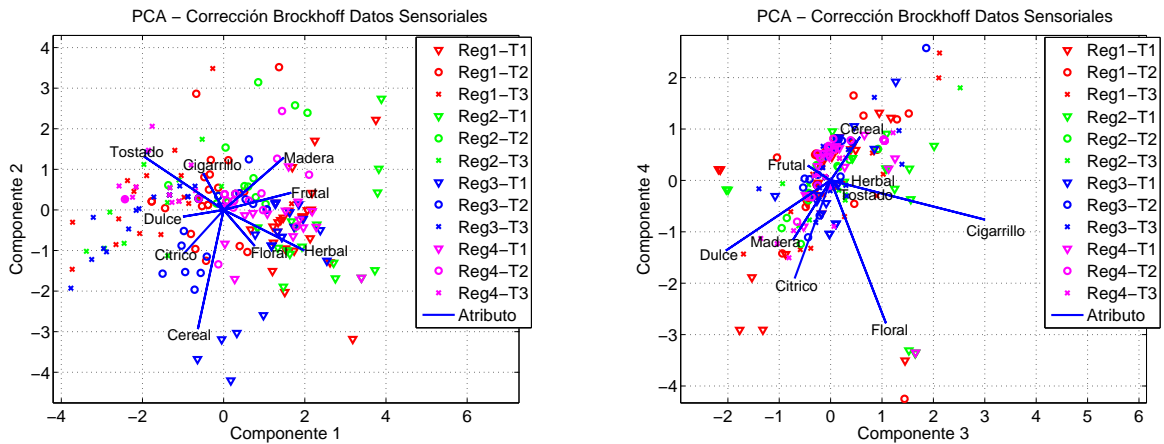


FIGURA 4.6. Biplot PCA con corrección Brockhoff para datos sensoriales (Escala Multiplicativa).

En la figura 4.7 se presentan los datos después de la transformación Brockhoff - escalado aditivo. Las regiones siguen mostrando traslape. De igual forma que en las transformadas anteriores, se presentan agrupaciones por temperaturas. En la gráfica de las componentes 3 y 4 no se pueden detectar agrupaciones.

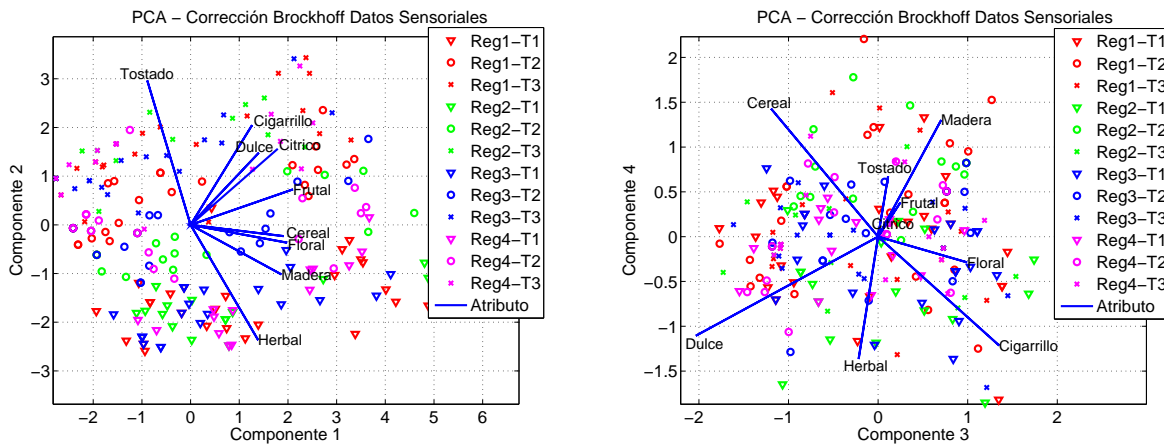


FIGURA 4.7. Biplot PCA con corrección Brockhoff para datos sensoriales (Escala Aditivo).

En la figura 4.8 se presenta nuevamente la transformación de Procrustes. En la gráfica se identifican las muestras que corresponden a las regiones usando los colores y los catadores usando los tipos de puntos. En estas gráficas se observa que la transformación eliminó el efecto del catador, lo cual concuerda con los resultados de ANOVA. Se pueden observar sub-agrupaciones correspondientes a una región para cierto catador, por ejemplo, se puede

notar que hay diferencias del catador dos para la región tres y región dos. Para las cuatro primeras PC se nota que el catador dos es el que muestra mayor dispersión.

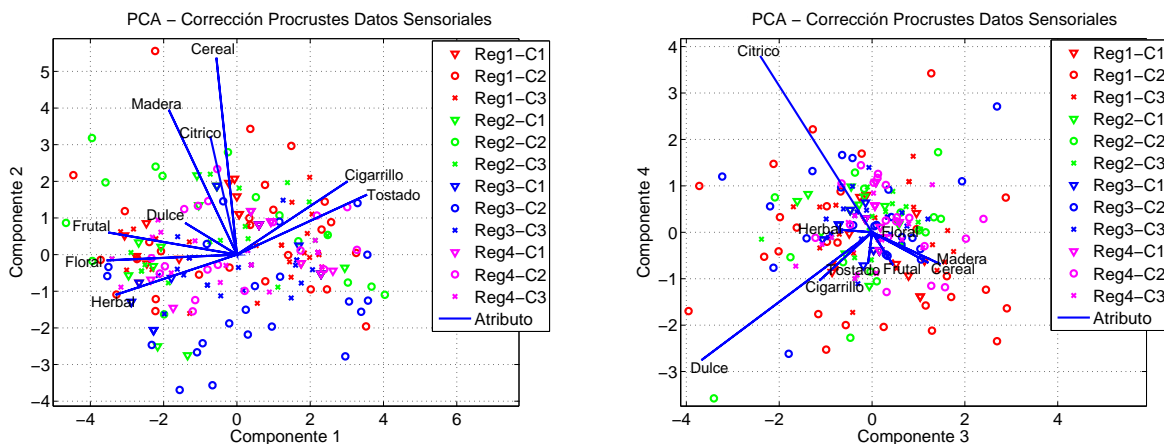


FIGURA 4.8. Biplot PCA con corrección Procrustes para datos sensoriales.

Después de la corrección de Brockhoff (Figura 4.9) se presenta el mismo fenómeno. Los catadores ya no forman agrupaciones pero se nota que el catador dos es el que presenta mayor dispersión. Nótese que, a pesar de la eliminación del efecto del catador, se podría diferenciar entre las regiones dos y tres considerando el catador dos.

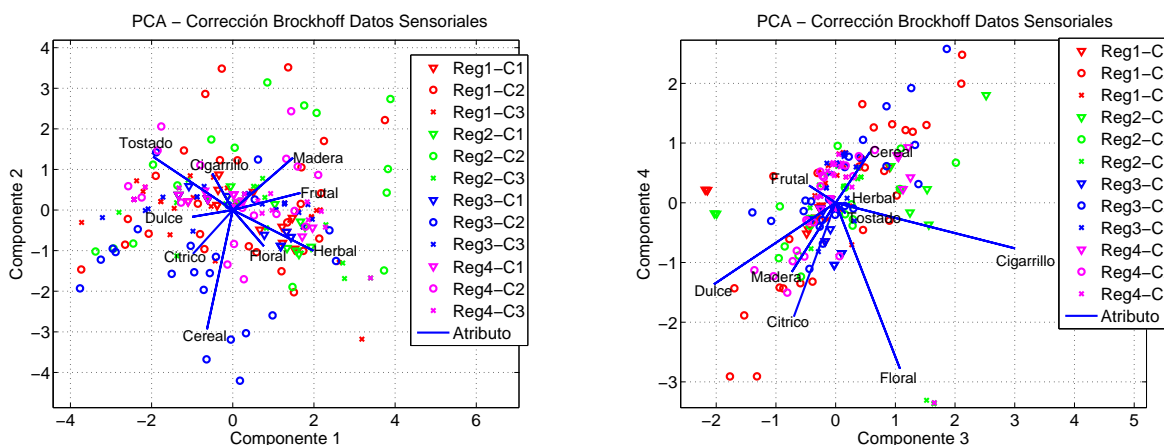


FIGURA 4.9. Biplot PCA Corrección Brockhoff Datos Sensoriales (Escala Multiplicativo).

El modelo del evaluador - escalado aditivo (ver figura 4.10) no afecta mucho las mediciones originales de los datos pues se siguen formando agrupaciones por catadores. A partir de las

cuatro primeras PC es difícil identificar agrupaciones por regiones para esta transformación.

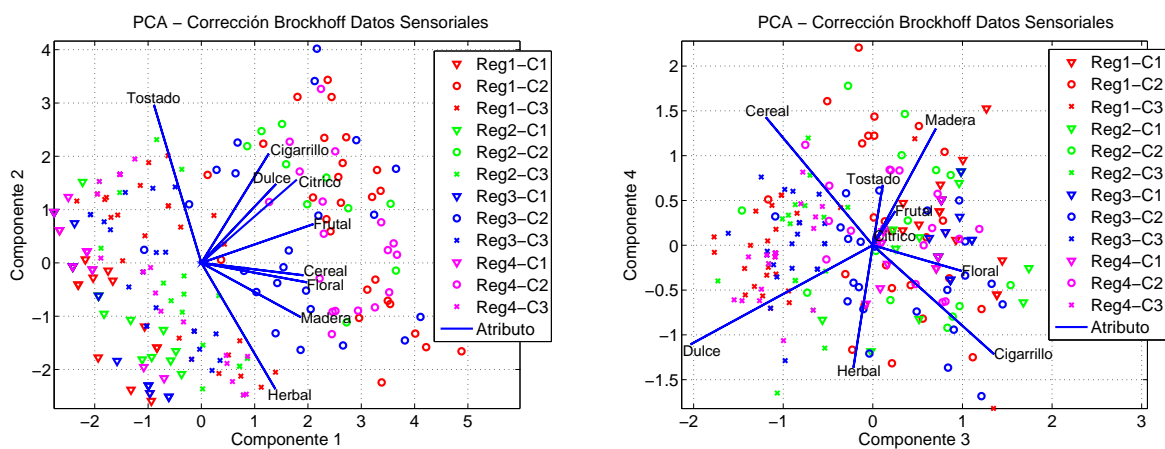


FIGURA 4.10. Biplot PCA Corrección Brockhoff Datos Sensoriales (Escala Aditivo).

Otro tipo de visualización de los factores del experimento es el que se forma de etiquetar temperaturas y catadores después de las transformaciones. En la figura 4.11 se observa el resultado de la transformación Procrustes al eliminar el efecto del catador. Se observa que independientemente del catador, las temperaturas son fáciles de identificar. Y se observa con mayor claridad que el catador dos es el que tiene mayor varianza. En la gráfica de los componentes tres y cuatro ya no es tan evidente las diferencias entre las regiones, pero se sigue presentando el mismo fenómeno de varianza en el segundo catador.

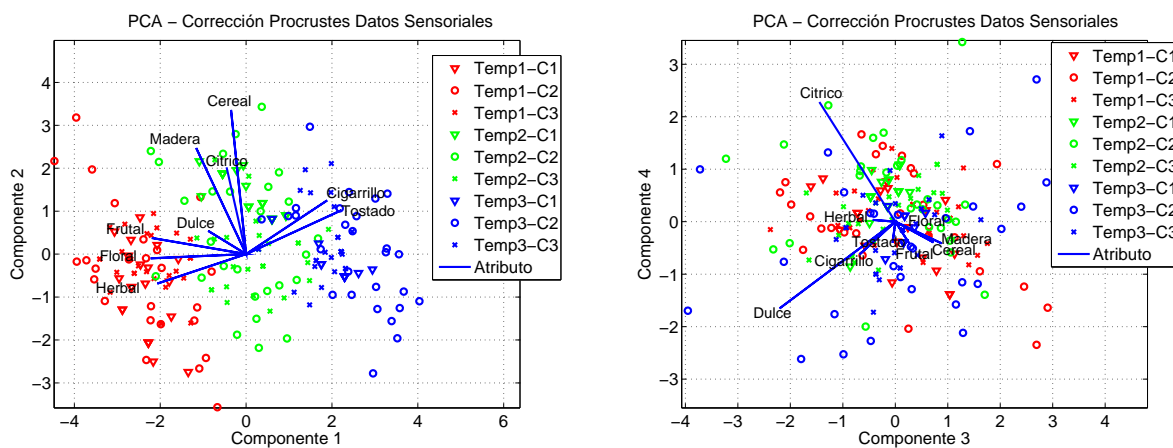


FIGURA 4.11. Biplot PCA Corrección Procrustes Datos Sensoriales .



Para la transformación con el modelo de evaluador de Brockhoff escalado multiplicativo (figura 4.12) se observan agrupaciones por temperaturas, pero también se observan unos valores atípicos después de la transformación. Lo anterior puede ser debido a que esta corrección al ser una transformación multiplicativa cambia drásticamente los datos.

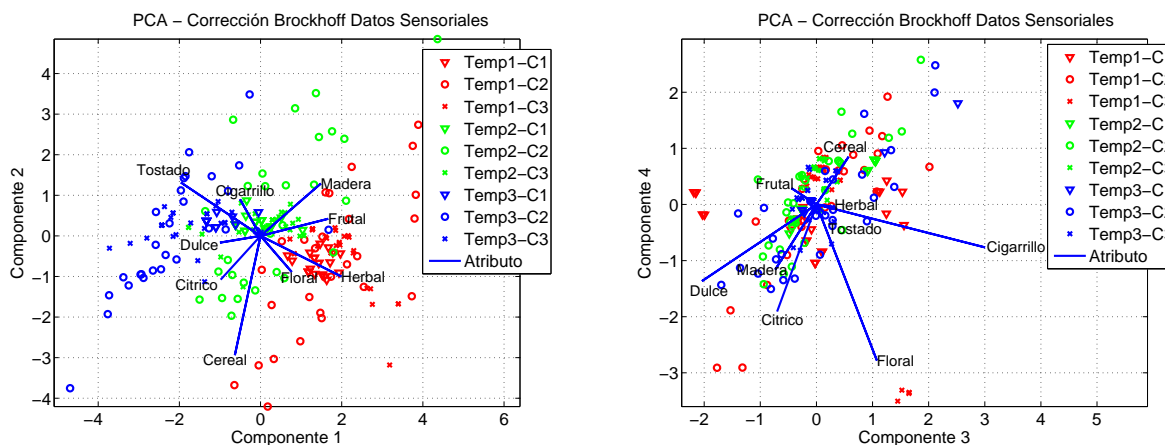


FIGURA 4.12. Biplot PCA Corrección Brockhoff Datos Sensoriales (Escala Multiplicativo).

Para la corrección de Brockhoff - escalado aditivo 4.13 sigue siendo evidente la diferencia entre las temperaturas, de igual forma que las diferencias entre catadores como para los datos sin corregir. Como lo demuestran los resultados de ANOVA, efecto del catador y la temperatura siguen siendo los efectos dominantes.

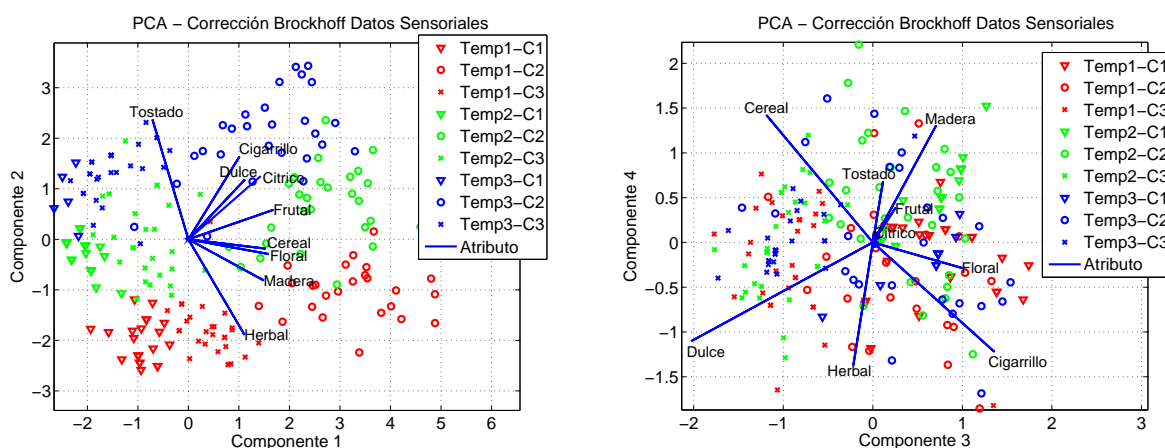


FIGURA 4.13. Biplot PCA Corrección Brockhoff Datos Sensoriales (Escala Aditivo).

#### 4.4. Desempeño de los clasificadores

Debido a la naturaleza multiclase de las bases de datos analizadas, es recomendable usar matrices de confusión para poder determinar cuántas observaciones de la clase  $i$  son clasificadas correctamente como de dicha clase o incorrectamente como de las clase  $j$ , para  $j \neq i$ .

La matriz de confusión de la figura 4.10 muestra los resultados de un problema de clasificación hipotético de 4 clases.

Clase real	Predicción de clase			
	clase 1	clase 2	...	clase n
clase 1	$a_{11}$	$a_{12}$	...	$a_{1n}$
clase 2	$a_{21}$	$a_{22}$	...	$a_{2n}$
⋮	⋮	⋮	⋮	⋮
clase n	$a_{n1}$	$a_{n2}$	...	$a_{nn}$

TABLA 4.10. Matriz de confusión

El cuerpo de la tabla tiene una fila y una columna por cada posible clasificación. Las filas corresponden a las etiquetas de clase verdaderas y las columnas corresponden a las clasificaciones predichas por el clasificador. [Bramer, 2007]

El valor de la  $i$ -ésima fila y  $j$ -ésima columna contiene el número de observaciones de la  $i$ -ésima clase que fueron clasificadas como de la  $j$ -ésima clase.

##### 4.4.1. Verdaderos y falsos positivos y negativos

Cuando una base de datos tiene solo dos clases, una se denomina "positiva" (o clase objetivo) y la otra se denomina "negativa". En este caso las entradas en la primera columna de la matriz de confusión se denominan *verdaderos* y *falsos positivos*, similarmente, los valores en la segunda columna se denominan *falsos* y *verdaderos negativos* (Tabla 4.11).

Clase real	Predicción de clase		Total de observaciones
	+	-	
+	verdaderos positivos ( $T_P$ )	falsos negativos ( $F_N$ )	Positivos ( $P$ )
-	falsos positivos ( $F_P$ )	verdaderos negativos ( $T_N$ )	Negativos ( $N$ )

TABLA 4.11. Verdaderos y falsos positivos y negativos

Cuando hay más de dos clases, y se quiere evaluar el desempeño de cada una de las clases en el problema de clasificación, se toma cada clase como positiva (una por vez), con todas las otras clases combinadas como clase negativa. Obteniendo una matriz de confusión por cada clase, lo cual simplifica el cálculo de las medidas de desempeño expuestas a continuación.

#### 4.4.2. Medidas de desempeño

Las medidas de desempeño empleadas típicamente se explican a continuación:

**Tasa de verdaderos positivos:** Es la proporción de las observaciones positivas que son clasificadas correctamente como positivas. También conocida como *Sensibilidad*.

$$T_{P_r} = T_P/P \quad (51)$$

**Tasa de falsos positivos:** Es la proporción de observaciones negativas que fueron clasificadas erróneamente como positivas.

$$F_{P_r} = F_P/N \quad (52)$$

**Tasa de falsos negativos:** Es la proporción de observaciones positivas que son clasificadas erróneamente como negativas.

$$F_{N_r} = F_N/P \quad F_{N_r} = 1 - T_{P_r} \quad (53)$$

**Tasa de verdaderos negativos:** Es la proporción de observaciones negativas que son correctamente clasificadas como negativas. También conocida como *Especificidad*.

$$T_{N_r} = T_N/N \quad T_{N_r} = 1 - F_{P_r} \quad (54)$$

**Precisión:** Es la proporción de observaciones que son clasificadas como positivas que son realmente positivas.

$$\text{Precisión} = T_P/(T_P + F_P) \quad (55)$$

Nótese que la Precisión solo penaliza los FP, pero aunque la proporción de TP sea baja la Precisión siempre daría 1 si no hay falsos positivos. Por lo tanto esta medida no se tuvo en cuenta para evaluar el desempeño general en los experimentos.

**Puntaje F1:** Es una medida que combina la Precisión y la tasa de verdaderos positivos.

$$F_1 = (2 \times \text{Precisión} \times T_{P_r})/(P + N) \quad (56)$$

Reemplazando Precisión por su equivalencia y simplificando se obtiene:

$$F_1 = \frac{2 \times \frac{T_P}{T_P + F_P} \times \frac{T_P}{P}}{\frac{T_P}{T_P + F_P} + \frac{T_P}{P}} = \frac{2 \times T_P}{T_P + F_P + P} \quad (57)$$

**Exactitud:** La proporción de observaciones que son clasificadas correctamente.

$$\text{Exactitud} = (T_P + T_N)/(P + N) \quad (58)$$

**Tasa de error:** Es la proporción de observaciones que son clasificadas incorrectamente.

$$\text{Error} = (F_N + F_P)/(P + N) \quad (59)$$

**Exactitud general:** Es la proporción de observaciones de cualquier clase que son clasificadas correctamente. Corresponde entonces a la traza de la matriz de confusión general mostrada en la tabla 4.10.

### 4.4.3. Validación cruzada

El método más simple y ampliamente usado para estimar el error de clasificación, garantizando su robustez estadística, es la validación cruzada [Bramer, 2007]. La validación cruzada del tipo  $K$ -fold usa una parte de los datos disponibles para ajustar el modelo (entrenamiento), y una parte diferente para validarlo. Los datos se dividen en  $K$  partes aproximadamente del mismo tamaño; por ejemplo, cuando  $K = 5$ , el conjunto de datos se dividiría como se muestra en la Figura 4.14:

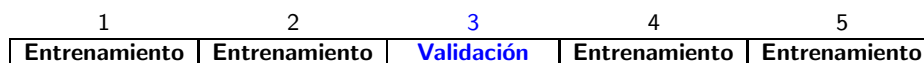


FIGURA 4.14. Particiones del conjunto de datos.

Para la  $k$ -ésima parte (tercera en el ejemplo de la figura 4.14), se entrena el clasificador con las restantes  $K - 1$  partes de los datos, y se calcula el error de clasificación usando los datos de la  $k$ -ésima parte como datos de validación. Lo anterior se realiza para  $k = 1, 2, \dots, K$ . Finalmente se promedian las  $K$  estimaciones del error de clasificación. El caso  $K = N$  se conoce como validación cruzada *leave-one-out*, siendo  $N$  el número total de muestras. Las elecciones típicas de  $K$  son 5 ó 10. Cuando  $K = 5$  la estimación del error tiene una varianza baja pero puede ser considerablemente sesgada. En [Bramer, 2007] se recomiendan el valor  $K = 10$  debido a que supone un compromiso aceptable entre el sesgo y la varianza de la estimación del error. Esta varianza se calcula como el error estándar de la estimación de la medida de desempeño, el cual está definido como  $Err_e = \sigma/\sqrt{n}$ . Donde  $\sigma$  es la desviación estándar del vector de estimaciones repetidas  $n$  veces, donde  $n$  es el número de repeticiones del experimento [Altman and Bland, 2005].

### 4.4.4. Comparación y evaluación mediante modelos discriminantes

A continuación se presentan los resultados de clasificación usando los clasificadores LDA y 1-NN, que fueron empleados por ser la alternativa más común para comparación de representaciones puesto que son los clasificadores más sencillos. Adicionalmente, se propone comparar los métodos de corrección de escala, entre ellos el modelo de evaluador propuesto por [Brockhoff and Skovgaard, 1994] y el método de Análisis de Procrustes con el método propuesto por [Granitto et al., 2008]. Éste último consiste en ajustar un clasificador individual para cada catador empleando los datos sin corrección. Posteriormente, se promedian las predicciones individuales para formar una predicción combinada. En esta etapa, se sigue el procedimiento típico en sistemas de clasificadores múltiples, se estima el promedio de la probabilidad a posteriori de todos los modelos discriminantes y se asigna a cada muestra el producto con máxima probabilidad a posteriori. Este método nos permitirá conocer cuál es la transformación más confiable de los datos o determinar si ninguna de éstas es mejor que promediar las predicciones de los modelos individuales.

En la tabla 4.12 se presentan las medidas de desempeño del clasificador LDA. Se consideraron probabilidades a priori iguales para los experimentos de regiones y los de temperaturas. De los resultados mostrados en la tabla se destacan las siguientes observaciones por región:

en el caso de la región 1, el mayor valor de verdaderos positivos corresponde a la combinación de clasificadores, con un 35.67%, seguido por la transformación de Brockhoff - escalado multiplicativo, con 30.28%, sin embargo este último presenta también la mayor tasa de falsos positivos, con una proporción de 21.08%, mientras que para la combinación de clasificadores los falsos positivos alcanzan una proporción de 14.11%. Las dos transformaciones: Brockhoff - escalado multiplicativo y Procrustes, y la estrategia de combinación de clasificación mejoran el desempeño obtenido con los datos sin corrección, mientras que la transformación de Brockhoff-escalado aditivo, deteriora ligeramente el desempeño para la región 1, y no aporta mejora alguna para las regiones 3 y 4.

Adicionalmente a las tasas de verdaderos y falsos positivos, se presentan otras dos medidas de desempeño, la exactitud y el puntaje F1, presentados en la sección 4.4.2. Puestos que estas dos medidas se pueden expresar en términos de los verdaderos positivos y falsos positivos, se espera que sean indicadores del desempeño global de clasificación para cada región. En particular para las regiones 1 y 2, se observan tasas bajas de verdaderos positivos y de falsos positivos, y esto implica que no se están asignando etiquetas de estas clases para observaciones de dichas clases o de alguna de las otras. Nótese que, a pesar de lo anterior, las tasas de exactitud son altas. Esta aparente incongruencia se explica mediante el siguiente análisis: la exactitud es una medida que combina los verdaderos positivos con los verdaderos negativos ( $1 - \text{falsos positivos}$ ) pero su confiabilidad es discutible en el caso multiclase, ya que no hay una ponderación equilibrada entre verdaderos positivos y verdaderos negativos. Dicho inequilibrio obedece a que al calcular la exactitud para cada clase, el problema queda desbalanceado con la clase negativa; por ejemplo, siendo esta última  $n - 1$  veces más grande que la clase positiva si  $n$  es el número de clases y si se considera el mismo número de ejemplos por clase. En el caso hipotético donde ninguna muestra es asignada a la clase positiva, la medida de exactitud puede alcanzar un valor de  $1 - (1/n)$  o más si las clases no tienen la misma cantidad de observaciones. Por esta razón, la medida de exactitud no permite hacer un análisis global del desempeño. Por otro lado, el puntaje F1 es una medida que combina la precisión con la sensibilidad; es decir, unifica los criterios de proporción de observaciones clasificadas como positivas que son realmente positivas con la proporción de observaciones positivas que fueron clasificadas correctamente como positivas (ver ecuaciones 56 y 57).

La transformación de Brockhoff – escalado multiplicativo presenta los mayores verdaderos positivos para la región 2, seguido por la transformación Procrustes; no obstante, ambos presentan también las más altas tasas de falsos positivos. A pesar de lo anterior, el puntaje F1 más alto corresponde a la combinación de clasificadores seguido por la transformada de Brockhoff – escalado multiplicativo. La identificación de esta región mejora para todos los métodos de corrección y la combinación de clasificadores con respecto a los datos sin corregir, sin embargo la transformación de Brockhoff (2) aporta únicamente un ligero incremento en la tasa de verdaderos positivos.

TABLA 4.12. Medidas de desempeño para clasificación de regiones y temperaturas usando LDA

Transformación de datos y estrategia de clasificación	Clases	Clasificador LDA			
		Verdaderos Positivos (%)	Falsos Positivos (%)	Exactitud (%)	Puntaje F1 (%)
Datos Originales	Región 1	16.11 ± 0.43	12.92 ± 0.23	63.43 ± 0.23	22.68 ± 0.55
Transformación Procrustes		24.67 ± 0.72	18.19 ± 0.52	62.76 ± 0.31	30.56 ± 0.68
Transformación Brockhoff <sup>a</sup>		30.28 ± 0.97	21.08 ± 0.66	62.70 ± 0.40	35.00 ± 0.84
Transformación Brockhoff (2) <sup>b</sup>		15.33 ± 0.36	13.03 ± 0.21	63.09 ± 0.20	21.68 ± 0.47
Combinación clasificadores		35.67 ± 1.21	14.11 ± 0.74	73.33 ± 0.52	39.99 ± 0.99
Datos Originales	Región 2	11.30 ± 0.61	13.26 ± 0.29	71.65 ± 0.26	13.71 ± 0.71
Transformación Procrustes		24.63 ± 0.94	19.31 ± 0.71	69.48 ± 0.46	24.30 ± 0.66
Transformación Brockhoff <sup>a</sup>		28.06 ± 0.92	19.65 ± 0.48	69.89 ± 0.39	27.10 ± 0.76
Transformación Brockhoff (2) <sup>b</sup>		13.24 ± 0.60	13.66 ± 0.26	71.72 ± 0.27	15.76 ± 0.70
Combinación clasif.		23.33 ± 1.08	13.83 ± 0.94	70.46 ± 0.78	28.34 ± 1.23
Datos Originales	Región 3	48.06 ± 0.38	36.26 ± 0.26	59.56 ± 0.21	38.79 ± 0.27
Transformación Procrustes		40.76 ± 0.42	21.97 ± 0.36	68.09 ± 0.25	40.52 ± 0.33
Transformación Brockhoff <sup>a</sup>		48.19 ± 1.17	32.35 ± 0.53	62.46 ± 0.47	40.59 ± 0.78
Transformación Brockhoff (2) <sup>b</sup>		48.06 ± 0.37	36.62 ± 0.22	59.30 ± 0.19	38.64 ± 0.26
Combinación clasif.		77.67 ± 1.21	19.33 ± 1.06	79.92 ± 0.83	66.07 ± 1.12
Datos Originales	Región 4	36.30 ± 0.53	33.10 ± 0.25	60.78 ± 0.18	27.00 ± 0.31
Transformación Procrustes		49.54 ± 1.20	27.99 ± 0.72	67.52 ± 0.55	37.88 ± 0.75
Transformación Brockhoff		24.26 ± 0.98	16.37 ± 0.52	71.76 ± 0.43	25.51 ± 0.89
Transformación Brockhoff (2)		36.30 ± 0.65	32.13 ± 0.17	61.56 ± 0.19	27.40 ± 0.43
Combinación clasif.		76.83 ± 1.29	14.89 ± 0.64	83.04 ± 0.66	69.41 ± 1.14
Datos Originales	Temp. 1	96.61 ± 0.05	3.69 ± 0.07	96.35 ± 0.05	94.64 ± 0.07
Transformación Procrustes		92.50 ± 0.81	3.94 ± 0.19	95.02 ± 0.21	92.58 ± 0.33
Transformación Brockhoff		95.28 ± 0.24	5.19 ± 0.16	89.89 ± 0.39	84.76 ± 0.58
Transformación Brockhoff (2)		95.83 ± 0.16	5.14 ± 0.14	96.98 ± 0.11	95.53 ± 0.15
Combinación clasif.		100.0 ± 0.00	0.38 ± 0.16	99.75 ± 0.12	99.63 ± 0.18
Datos Originales	Temp. 2	85.94 ± 0.15	6.47 ± 0.08	90.93 ± 0.09	86.27 ± 0.15
Transformación Procrustes		82.78 ± 0.63	9.28 ± 0.50	88.59 ± 0.29	82.77 ± 0.39
Transformación Brockhoff		79.33 ± 0.37	7.56 ± 0.16	80.98 ± 0.46	72.28 ± 0.50
Transformación Brockhoff (2)		79.22 ± 0.40	7.44 ± 0.16	91.44 ± 0.18	87.16 ± 0.28
Combinación clasif.		97.00 ± 0.59	0.50 ± 0.14	98.75 ± 0.23	98.09 ± 0.36
Datos Originales	Temp. 3	90.44 ± 0.15	3.33 ± 0.00	94.57 ± 0.07	91.76 ± 0.11
Transformación Procrustes		88.94 ± 0.69	4.67 ± 0.20	93.57 ± 0.24	90.32 ± 0.40
Transformación Brockhoff		88.56 ± 0.31	5.67 ± 0.12	90.39 ± 0.24	85.22 ± 0.39
Transformación Brockhoff (2)		88.28 ± 0.28	5.75 ± 0.11	94.46 ± 0.11	91.59 ± 0.17
Combinación clasif.		99.00 ± 0.28	1.13 ± 0.25	99.00 ± 0.20	98.52 ± 0.29

a. Brockhoff – escalado multiplicativo

b. Brockhoff – escalado aditivo

Las tasas de verdaderos positivos para la región 3 son las más altas en comparación con las otras regiones. No obstante, los falsos positivos, para dicha región, también presentan algunas de las más altas tasas. Lo anterior indica que este clasificador está dominado por la región 3, pues una gran cantidad de observaciones pertenecientes a las otras clases están siendo asignadas a esta región. La mayor tasa de verdaderos positivos para la región 3 la

proporciona la combinación de clasificadores, dando a la vez la menor tasa de falsos positivos. Por consiguiente, sus medidas de exactitud y puntaje F1, son las más altas para esta región. Las transformaciones de Brockhoff, para los dos tipos de escalado, tienen tasas de verdaderos positivos y falsos positivos casi iguales a las obtenidas con los datos sin corregir. Nótese que la tasa de verdaderos positivos de la transformación Procrustes no es mejor que la de los datos sin corregir, aunque su tasa de falsos positivos sí es menor. Con excepción del puntaje F1 para la combinación de clasificadores, esta medida es similar para las transformaciones y los datos sin corregir.

Similarmente como ocurre con la región 3 se presentan altas tasas de verdaderos positivos y falsos positivos para la región 4. También, para el caso de la combinación de clasificadores se presenta el mejor compromiso entre las dos medidas anteriores como lo sugiere el puntaje F1.

El otro subproblema de clasificación, que tiene que ver con la clasificación de las temperaturas resulta ser un problema más sencillo para el clasificador lineal. En la tabla 4.12 puede observarse que, consistentemente con el problema anterior la combinación de clasificadores tiene de nuevo los mejores desempeño respecto a todas las medidas. En este caso, a diferencia de la clasificación de regiones, las transformaciones de Brockhoff y Procrustes no contribuyen a mejorar el desempeño sino que, por el contrario, deterioran los desempeños obtenidos con los datos originales. Siendo el clasificador empleado de frontera de decisión lineal, se deduce que las transformaciones si mejoran la separabilidad, dicha condición es no lineal.

Teniendo en cuenta que la estrategia de combinación de clasificadores proporciona los mejores resultados para la identificación y caracterización de las regiones y temperaturas, es interesante conocer el desempeño individual de los catadores. Para identificar las habilidades individuales, a continuación se discuten las matrices de confusión de los resultados de clasificación, previos a la combinación, para cada catador. En la tabla 4.13 se presentan las matrices de confusión de los tres catadores, para el problema de la identificación de regiones. Para la región 1 los catadores 1 y 2 son los que presentan mayores aciertos, no obstante las mayores confusiones de esta clase consisten en la asignación de sus observaciones para las regiones 3 y 4. Se observa en estas matrices que la región con más asignaciones es la región 3, de tal forma que los tres catadores identifican bien esta región. De manera opuesta, ninguno de los catadores logra identificar una cantidad considerable de observaciones de la región 2. Finalmente, el catador 1 es el mejor para identificar la región 4.

TABLA 4.13. Matrices de confusión de los catadores usando LDA para identificación de las regiones

(a) Catador 1						(b) Catador 2						(c) Catador 3								
		Clase predicha				Total			Clase predicha				Total			Clase predicha				Total
Clase real		<b>13</b>	2	<b>14</b>	1	30	Clase real		5	4	<b>12</b>	<b>9</b>	30	Clase real		<b>10</b>	5	<b>7</b>	<b>8</b>	30
		4	3	6	5	18			5	3	4	6	18			7	6	4	1	18
		3	4	<b>17</b>	0	24			3	4	<b>16</b>	1	24			3	4	<b>14</b>	3	24
		0	2	5	<b>11</b>	18			5	5	2	6	18			5	4	4	5	18
Total		20	11	<b>42</b>	17	90	Total		18	16	<b>34</b>	22	90	Total		25	19	<b>29</b>	17	90

En la tabla 4.14 se presentan los resultados de clasificación usando el clasificador del vecino más cercano para los problemas de identificación de regiones y de temperaturas. Para el primer problema, si bien para la región 1 la mejor tasa de verdaderos positivos es la de la combinación de clasificadores, la tasa de falsos positivos es alta. Para esta región, un mejor compromiso entre verdaderos positivos y falsos positivos se obtiene con la transformación de Brockhoff (2) - escalado aditivo, con un puntaje F1 de 60.66%, mientras dicho puntaje para la combinación es de 55.43%. se observa que la región 2 sí se puede separar de las demás regiones por medio del clasificador 1-NN que no se podía identificar con el clasificador LDA. Por consiguiente, existe una frontera natural no lineal entre esta región y las demás. El mejor desempeño se obtiene de la combinación de clasificadores. Nótese además que el segundo mejor desempeño corresponde a la transformación de Brockhoff (2). Más aún, para las otras regiones estos dos clasificadores son los que mejor desempeño tienen, con una diferencia de 3.87% para la región 3 y apenas un 1.31% para la región 4.

En el caso de discriminación de temperaturas, así como ocurrió para el clasificador LDA, el mejor desempeño lo presenta la combinación de clasificadores. Además, como para los resultados en la tabla 4.12, véase en 4.14 que ninguna transformación mejora el desempeño de clasificación de los datos originales, excepto por una ligera ventaja de la transformación Procrustes en la identificación de la temperatura 3.

En resumen, el problema de identificación de temperaturas es linealmente separable. La superioridad del clasificador LDA frente al clasificador 1-NN en este problema debe obedecer a una mejor capacidad de generalización. Es probable que el clasificador 1-NN se sobreajuste a las no linealidades de las fronteras en el conjunto de entrenamiento, mientras que la frontera del clasificador LDA no se ve distorsionada por la presencia de unos cuantos objetos que invadan el territorio de otras clases. Para el caso de la identificación de regiones, el comportamiento del problema es precisamente opuesto. Los mejores desempeños corresponden al clasificador 1-NN, lo cual revela la naturaleza no lineal de la frontera entre las clases.



TABLA 4.14. Medidas de desempeño para clasificación de regiones y temperaturas usando 1-NN

Transformación de Datos	Clases	Clasificador 1-NN			
		Verdaderos Posit. (%)	Falsos Positivos (%)	Exactitud (%)	Puntaje F1 (%)
Datos sin corregir	Región 1	64.67 ± 0.34	41.06 ± 0.25	60.24 ± 0.24	51.85 ± 0.26
Transformación Procrustes		62.89 ± 0.93	33.94 ± 0.41	65.57 ± 0.58	55.51 ± 0.69
Transformación Brockhoff		62.61 ± 0.41	37.56 ± 0.30	63.26 ± 0.31	53.76 ± 0.41
Transformación Brockhoff (2)		59.39 ± 0.50	19.47 ± 0.42	73.81 ± 0.26	60.66 ± 0.31
Combinación clasificadores		77.50 ± 1.32	35.89 ± 0.78	68.75 ± 0.63	55.43 ± 0.76
Datos sin corregir	Región 2	49.91 ± 0.28	10.00 ± 0.23	81.94 ± 0.19	52.53 ± 0.36
Transformación Procrustes		36.94 ± 0.89	18.54 ± 0.63	73.52 ± 0.45	35.67 ± 1.11
Transformación Brockhoff		46.30 ± 0.71	10.02 ± 0.29	81.19 ± 0.23	50.46 ± 0.64
Transformación Brockhoff (2)		59.44 ± 0.68	11.25 ± 0.22	83.07 ± 0.31	58.00 ± 0.69
Combinación clasificadores		61.00 ± 1.59	7.44 ± 0.45	85.75 ± 0.61	68.79 ± 1.36
Datos sin corregir	Región 3	37.29 ± 0.25	9.62 ± 0.30	76.54 ± 0.17	45.70 ± 0.31
Transformación Procrustes		32.01 ± 0.68	8.81 ± 0.32	74.76 ± 0.40	40.27 ± 0.72
Transformación Brockhoff		39.24 ± 0.53	14.22 ± 0.30	73.94 ± 0.21	45.08 ± 0.49
Transformación Brockhoff (2)		60.35 ± 0.68	10.45 ± 0.37	80.94 ± 0.24	62.37 ± 0.46
Combinación clasificadores		47.00 ± 1.43	5.17 ± 0.47	82.79 ± 0.49	58.50 ± 1.21
Datos sin corregir	Región 4	35.56 ± 0.44	11.23 ± 0.26	78.46 ± 0.22	40.50 ± 0.60
Transformación Procrustes		41.02 ± 0.76	13.73 ± 0.42	76.89 ± 0.35	40.91 ± 0.87
Transformación Brockhoff		36.48 ± 0.41	10.79 ± 0.28	78.80 ± 0.23	39.52 ± 0.38
Transformación Brockhoff (2)		55.28 ± 0.69	14.40 ± 0.32	80.09 ± 0.29	53.16 ± 0.55
Combinación clasificadores		44.33 ± 1.41	8.22 ± 0.49	80.62 ± 0.40	54.47 ± 0.10
Datos sin corregir	Temp. 1	95.67 ± 0.24	3.22 ± 0.11	96.30 ± 0.12	94.51 ± 0.18
Transformación Procrustes		89.94 ± 0.74	5.56 ± 0.23	93.26 ± 0.25	89.90 ± 0.40
Transformación Brockhoff		93.39 ± 0.37	3.39 ± 0.13	93.98 ± 0.25	91.14 ± 0.37
Transformación Brockhoff (2)		93.44 ± 0.33	3.11 ± 0.16	96.15 ± 0.15	94.29 ± 0.22
Combinación clasificadores		99.88 ± 0.64	0.13 ± 0.13	99.71 ± 0.15	99.56 ± 0.23
Datos sin corregir	Temp. 2	86.33 ± 0.23	11.28 ± 0.19	87.56 ± 0.18	82.03 ± 0.26
Transformación Procrustes		80.89 ± 0.57	11.53 ± 0.47	85.74 ± 0.47	79.09 ± 0.62
Transformación Brockhoff		84.89 ± 0.23	13.33 ± 0.21	83.96 ± 0.33	76.49 ± 0.46
Transformación Brockhoff (2)		84.94 ± 0.46	13.25 ± 0.20	89.74 ± 0.19	84.62 ± 0.28
Combinación clasificadores		97.63 ± 0.59	4.44 ± 0.31	96.21 ± 0.35	94.50 ± 0.50
Datos sin corregir	Temp. 3	81.78 ± 0.31	3.61 ± 0.12	91.26 ± 0.13	86.18 ± 0.20
Transformación Procrustes		86.33 ± 0.62	4.33 ± 0.28	92.41 ± 0.29	88.29 ± 0.47
Transformación Brockhoff		79.17 ± 0.31	4.56 ± 0.16	89.72 ± 0.18	83.89 ± 0.31
Transformación Brockhoff (2)		79.44 ± 0.34	4.72 ± 0.17	93.59 ± 0.20	90.27 ± 0.32
Combinación clasificadores		91.25 ± 0.13	1.06 ± 0.23	96.50 ± 0.32	94.56 ± 0.52

a. Brockhoff – escalado multiplicativo

b. Brockhoff – escalado aditivo

Nótese que la exactitud global (tabla 4.16) no corresponde al promedio de las exactitudes por clase, sino que es el porcentaje total de etiquetas bien asignadas. La combinación de clasificadores siempre produce buenos resultados independientemente del clasificador empleado, en cambio la transformación de Brockhoff (2) produce buenos resultados únicamente para el clasificador 1-NN, siendo por lo tanto dependiente de la naturaleza del método de clasificación.

TABLA 4.15. Matrices de confusión de los catadores usando 1-NN para identificación de las regiones

(a) Catador 1						(b) Catador 2						(c) Catador 3								
		Clase predicha				Total			Clase predicha				Total			Clase predicha				Total
Clase real		27	2	1	0	30	Clase real		15	5	5	5	30	Clase real		18	4	4	4	30
		10	4	0	4	18			3	12	1	2	18			4	10	0	4	18
		16	0	8	0	24			9	1	10	4	24			13	1	8	2	24
		7	3	0	8	18			8	2	0	8	18			9	4	2	3	18
Total		60	9	9	12	90	Total		35	20	16	19	90	Total		44	19	14	13	90

TABLA 4.16. Medidas de exactitud global para los problemas de regiones y temperaturas

Transformación de Datos	Clases	Clasificador	
		Exactitud LDA(%)	Exactitud 1-NN(%)
Datos sin corregir	Regiones	$27.70 \pm 0.27$	$48.59 \pm 0.21$
Transformación Procrustes		$33.93 \pm 0.36$	$45.37 \pm 0.47$
Transformación Brockhoff		$33.41 \pm 0.45$	$48.59 \pm 0.26$
Transformación Brockhoff (2)		$27.83 \pm 0.24$	$58.96 \pm 0.30$
Combinación clasificadores		$53.38 \pm 0.55$	$59.96 \pm 0.72$
Datos sin corregir		Temperaturas	$90.93 \pm 0.09$
Transformación Procrustes	$88.59 \pm 0.29$		$85.70 \pm 0.47$
Transformación Brockhoff	$80.63 \pm 0.38$		$83.83 \pm 0.34$
Transformación Brockhoff (2)	$91.44 \pm 0.18$		$89.74 \pm 0.19$
Combinación clasificadores	$98.75 \pm 0.23$		$96.21 \pm 0.35$

a. Brockhoff – escalado multiplicativo

b. Brockhoff – escalado aditivo

#### 4.4.5. Resultados para la base de datos de papas

A continuación se presentan los resultados de los clasificadores LDA y 1-NN sobre la base de datos de papa. Se discuten y comparan las tasas de desempeño sobre los datos después de aplicar todas las transformaciones que eliminan el efecto del catador junto con la estrategia de combinación de clasificadores, con el fin de determinar si las transformaciones permiten mejorar la separabilidad entre clases o si es conveniente aprovechar la experticia individual.

En la tabla 4.17 se presentan los resultados de las tasas de verdaderos positivos, falsos positivos, precisión y puntaje F1 para el clasificador LDA, calculados a partir de los datos sin corregir, con correcciones y usando la combinación de clasificadores, para cada una de las clases del problema, i.e., seis variedades de papa.

Al emplear el clasificador LDA, se destaca que la variedad 2 constituye la clase más definida del problema, puesto que posee los valores de verdaderos positivos más altos y los falsos positivos más bajos, y consecuentemente el mayor puntaje F1. Particularmente, la estrategia de combinación de clasificadores es la que proporciona el mejor resultado para la variedad 2. De hecho, excepto por el caso de la variedad 5, la combinación de clasificadores representa la mejor alternativa para la identificación de variedades. Nótese que en general, las transformaciones no mejoran la separabilidad, excepto para el caso de la transformación

de Brockhoff (2), que presenta una mejoría en la precisión para las variedades 1, 5 y 6. Además, se intuye un severo traslape de clases, pues todos los desempeños son pobres.

TABLA 4.17. Desempeño de clasificación para el problema de papa usando el clasificador LDA sobre los datos

Transformación de Datos	Clases	Clasificador LDA			
		Verdaderos Posit. (%)	Falsos Positivos (%)	Precisión(%)	Puntaje F1 (%)
Datos sin corregir	Variedad 1	34,79 ± 0,99	10,98 ± 0,23	79,98 ± 0,27	36,63 ± 0,90
Transformación Procrustes		16,15 ± 1,08	8,50 ± 0,28	78,94 ± 0,27	20,19 ± 1,23
Transformación Brockhoff		28,75 ± 0,94	10,46 ± 0,18	79,41 ± 0,22	31,69 ± 0,90
Transformación Brockhoff (2)		37,60 ± 0,78	15,06 ± 0,25	77,05 ± 0,28	35,32 ± 0,70
Combinación clasificadores		36,25 ± 2,85	5,37 ± 0,55	84,89 ± 0,58	43,58 ± 2,72
2 Datos sin corregir		Variedad 2	83,13 ± 0,14	10,54 ± 0,14	88,40 ± 0,11
Transformación Procrustes	39,58 ± 0,79		12,63 ± 0,31	79,41 ± 0,27	39,04 ± 0,67
Transformación Brockhoff	86,35 ± 0,58		8,23 ± 0,21	90,87 ± 0,19	75,93 ± 0,44
Transformación Brockhoff (2)	30,00 ± 0,59		0,85 ± 0,02	87,62 ± 0,10	44,63 ± 0,69
Combinación clasificadores	99,37 ± 0,62		6,87 ± 0,60	94,16 ± 0,51	85,29 ± 1,22
Datos sin corregir	Variedad 3		64,17 ± 0,54	10,10 ± 0,13	85,61 ± 0,13
Transformación Procrustes		25,52 ± 0,83	14,17 ± 0,36	75,78 ± 0,28	25,94 ± 0,68
Transformación Brockhoff		63,85 ± 1,12	12,40 ± 0,16	83,65 ± 0,26	56,50 ± 0,82
Transformación Brockhoff (2)		27,29 ± 0,60	3,79 ± 0,10	84,72 ± 0,14	37,28 ± 0,69
Combinación clasificadores		76,87 ± 1,37	14,62 ± 0,93	83,96 ± 0,87	61,89 ± 1,64
Datos Originales		Variedad 4	51,98 ± 0,65	11,52 ± 0,20	82,40 ± 0,18
Transformación Procrustes	25,31 ± 1,37		18,17 ± 0,44	72,41 ± 0,34	23,26 ± 1,06
Transformación Brockhoff	46,46 ± 0,43		12,00 ± 0,26	81,08 ± 0,24	45,03 ± 0,47
Transformación Brockhoff (2)	47,29 ± 0,57		13,29 ± 0,17	80,14 ± 0,17	44,24 ± 0,45
Combinación clasificadores	53,75 ± 2,23		10,12 ± 0,53	83,85 ± 0,71	52,55 ± 2,07
Datos sin corregir	Variedad 5		26,88 ± 0,58	9,88 ± 0,24	79,58 ± 0,22
Transformación Procrustes		24,58 ± 0,73	22,19 ± 0,52	68,94 ± 0,46	20,89 ± 0,61
Transformación Brockhoff		26,67 ± 0,84	10,06 ± 0,26	79,39 ± 0,24	30,08 ± 0,81
Transformación Brockhoff (2)		48,13 ± 0,67	20,73 ± 0,18	74,08 ± 0,20	38,22 ± 0,47
Combinación clasificadores		19,37 ± 2,93	5,00 ± 0,51	82,39 ± 0,49	25,30 ± 3,36
Datos sin corregir		Variedad 6	19,58 ± 0,67	10,88 ± 0,18	77,53 ± 0,21
Transformación Procrustes	20,31 ± 0,85		14,06 ± 0,36	75,00 ± 0,30	21,25 ± 0,77
Transformación Brockhoff	21,35 ± 0,78		12,17 ± 0,39	76,75 ± 0,37	23,45 ± 0,87
Transformación Brockhoff (2)	32,50 ± 0,72		21,71 ± 0,22	70,66 ± 0,20	26,94 ± 0,50
Combinación clasificadores	54,37 ± 2,61		10,00 ± 0,89	84,06 ± 0,92	53,31 ± 2,40

a. Brockhoff – escalado multiplicativo

b. Brockhoff – escalado aditivo

Con el fin de ilustrar las confusiones entre clases, se indican a continuación las confusiones más importantes, de aquellos valores porcentuales mostradas en la matriz de confusión de la tabla 4.18: La variedad 1 se confunde en proporciones similares con las clases 5 y 6, casi la mitad de las observaciones de la variedad 1 son asignadas a estas dos clases. La confusión entre la variedades 1 y 6 es recíproca, puesto que se presenta una gran confusión de observaciones de la variedad 6 asignadas como de la variedad 1. Es notable que la variedad 4 genera altas confusiones con las variedades 3, 5 y 6. Lo anterior sugiere que la variedad 4 tiene una alta dispersión en el espacio de representación, que se superpone significativamente con el territorio de las variedades mencionadas.

TABLA 4.18. Matriz de confusión porcentual de la combinación de clasificadores usando el clasificador LDA

		Clase predicha					
		40.62	9.37	0	4.37	23.12	22.50
Clase real	0.62	95.00	0	3.12	0	1.25	
	0	0	80.62	18.12	1.25	0	
	6.25	8.12	35.62	13.12	15.00	21.87	
	10.62	0	11.25	18.12	58.12	1.87	
	36.87	6.25	11.87	15.00	5.00	25.00	

TABLA 4.19. Desempeño de clasificación para el problema de papa usando el clasificador 1-NN sobre los datos

Transformación de Datos	Clases	Clasificador 1-NN			
		Verdaderos Posit. (%)	Falsos Positivos (%)	Precisión(%)	Puntaje F1 (%)
Datos sin corrección	Variedad 1	29,27 ± 0,72	10,98 ± 0,18	77,69 ± 0,17	30,39 ± 0,60
Transformación Procrustes		28,54 ± 1,18	8,50 ± 0,62	77,10 ± 0,51	29,33 ± 1,06
Transformación Brockhoff		33,75 ± 1,04	10,46 ± 0,32	77,64 ± 0,35	33,45 ± 0,97
Transformación Brockhoff (2)		15,21 ± 0,59	15,06 ± 0,36	73,06 ± 0,31	15,82 ± 0,58
Combinación clasificadores		34,38 ± 3,25	11,37 ± 0,64	79,58 ± 0,60	34,98 ± 2,56
2 Datos sin corrección	Variedad 2	72,29 ± 0,55	10,54 ± 0,19	90,83 ± 0,18	72,45 ± 0,48
Transformación Procrustes		25,10 ± 1,23	12,63 ± 0,30	78,30 ± 0,32	27,71 ± 1,22
Transformación Brockhoff		67,19 ± 0,99	8,23 ± 0,31	89,10 ± 0,31	67,27 ± 0,85
Transformación Brockhoff (2)		8,96 ± 0,53	0,85 ± 0,07	84,05 ± 0,11	15,70 ± 0,86
Combinación clasificadores		93,13 ± 2,12	5,13 ± 0,34	94,58 ± 0,55	85,05 ± 1,60
Datos sin corrección	Variedad 3	48,13 ± 0,62	10,10 ± 0,24	81,08 ± 0,25	45,89 ± 0,59
Transformación Procrustes		31,25 ± 0,97	14,17 ± 0,41	77,20 ± 0,39	31,36 ± 0,90
Transformación Brockhoff		43,75 ± 0,54	12,40 ± 0,33	80,43 ± 0,26	42,72 ± 0,46
Transformación Brockhoff (2)		18,75 ± 1,09	3,79 ± 0,18	80,61 ± 0,26	24,25 ± 1,28
Combinación clasificadores		78,75 ± 2,42	11,88 ± 0,47	86,56 ± 0,38	65,92 ± 1,22
Datos sin corrección	Variedad 4	34,17 ± 0,70	11,52 ± 0,25	77,90 ± 0,29	34,02 ± 0,72
Transformación Procrustes		32,08 ± 1,40	18,17 ± 0,55	72,45 ± 0,43	27,85 ± 1,00
Transformación Brockhoff		35,42 ± 0,80	12,00 ± 0,29	77,73 ± 0,26	34,62 ± 0,66
Transformación Brockhoff (2)		33,54 ± 0,74	13,29 ± 0,34	72,83 ± 0,30	29,15 ± 0,58
Combinación clasificadores		08,13 ± 2,08	12,63 ± 0,66	74,16 ± 0,65	9,00 ± 2,30
Datos sin corrección	Variedad 5	28,96 ± 0,52	9,88 ± 0,25	77,01 ± 0,22	29,57 ± 0,47
Transformación Procrustes		26,04 ± 1,35	22,19 ± 0,67	75,16 ± 0,62	25,90 ± 1,32
Transformación Brockhoff		18,23 ± 0,85	10,06 ± 0,38	73,99 ± 0,31	18,87 ± 0,73
Transformación Brockhoff (2)		48,02 ± 1,00	20,73 ± 0,42	61,65 ± 0,43	29,46 ± 0,61
Combinación clasificadores		55,63 ± 1,92	10,00 ± 1,00	84,27 ± 0,90	54,44 ± 1,87
Datos sin corrección	Variedad 6	25,31 ± 0,63	10,88 ± 0,20	74,86 ± 0,20	25,11 ± 0,57
Transformación Procrustes		16,15 ± 1,15	14,06 ± 0,67	72,85 ± 0,51	16,38 ± 1,01
Transformación Brockhoff		22,08 ± 0,76	12,17 ± 0,30	74,58 ± 0,29	22,44 ± 0,74
Transformación Brockhoff (2)		13,54 ± 0,63	21,71 ± 0,41	73,82 ± 0,38	14,73 ± 0,69
Combinación clasificadores		23,13 ± 2,08	10,38 ± 0,69	78,54 ± 0,62	26,03 ± 2,17

a. Brockhoff – escalado multiplicativo

b. Brockhoff – escalado aditivo

En la tabla 4.19 se presentan las medidas de desempeño del clasificador 1-NN. Se pueden destacar las siguientes observaciones y comparaciones de los dos clasificadores empleados:

las regiones 6 y 4 se generalizan mejor a través de una función lineal, mientras que ellas mismas discriminadas con el clasificador knn sufren de sobreajuste con el clasificador. La variedad 2 sigue siendo la más fácil de discriminar, como ocurre con el clasificador LDA con un puntaje F1 de 85.05% usando la combinación de clasificadores. La combinación de clasificadores usando 1-NN sigue siendo la mejor alternativa para la discriminación de variedades de papa. Las únicas excepciones a este comportamiento general corresponden a la variedad 4 usando 1-NN. Se observa que la remoción del efecto de los catadores usando las transformaciones no mejora el desempeño, sino que, por el contrario, lo deteriora. Lo anterior indica que las diferencias entre catadores resultan relevantes en la discriminación de las variedades.

TABLA 4.20. Medidas de exactitud global para el problema de variedades de papa

Transformación de Datos	Clasificador	
	Exactitud LDA(%)	Exactitud 1-NN(%)
Datos sin corregir	46.75 ± 0.25	39.69 ± 0.24
Transformación Procrustes	25.24 ± 0.33	26.53 ± 0.64
Transformación Brockhoff	45.57 ± 0.40	36.74 ± 0.24
Transformación Brockhoff (2)	37.14 ± 0.25	23.00 ± 0.42
Combinación clasificadores	55.42 ± 1.05	52.08 ± 1.05

a. Brockhoff – escalado multiplicativo

b. Brockhoff – escalado aditivo

En la tabla 4.20 se presentan los resultados de exactitudes globales para los dos clasificadores empleados. Nótese que el mejor resultado global corresponde a la combinación de clasificadores con clasificador lineal con un valor de 55.42%. Este problema de identificación de variedades de papa es complejo, pues en general las tasas de exactitud son muy bajas.

#### 4.4.6. Resultados de la base de datos de queso

Similarmente al problema anterior, la identificación de tipos de queso es también un problema complejo, no obstante, en este caso el mejor desempeño el logrado para los datos sin corregir y el clasificador lineal. El pobre desempeño de la estrategia de combinación de clasificadores se explica en la no disponibilidad de una muestra de tamaño suficiente, puesto que se cuenta en este conjunto de datos con solo tres repeticiones por catador, dificultándose enormemente la generación de un modelo para cada catador con 10 clases y 3 muestras por clase.

TABLA 4.21. Medidas de exactitud global para el problema de tipos de queso

Transformación de Datos	Clasificador	
	Exactitud LDA(%)	Exactitud 1-NN(%)
Datos sin corregir	59.87 ± 0.48	51.13 ± 0.55
Transformación Procrustes	22.12 ± 0.50	23.71 ± 0.36
Transformación Brockhoff	54.65 ± 0.47	44.02 ± 0.56
Transformación Brockhoff (2)	47.90 ± 0.41	32.52 ± 0.42
Combinación clasificadores	34.17 ± 1.87	51.83 ± 1.40

## 4.5. Preprocesamiento de señales de cromatografía

Como se mencionó en el marco teórico (sección 3.2), es común obtener señales de cromatografía con cierto desplazamiento entre ellas, además de diferencias de escala en el eje del tiempo. Con el fin de poder emplear análisis multivariado sobre las señales de cromatografía, es necesario lograr hacer coincidir, hasta donde sea posible, los puntos correspondientes entre las señales, tal que puedan ser interpretados como una misma características en todas ellas.

### 4.5.1. Alineamiento de cromatografías

Las señales de cromatografía obtenidas para la base de datos de café fueron tomadas en dos periodos de tiempo diferentes, a saber, la mitad de ellas durante 3 semanas del mes de junio de 2009 y la otra mitad durante 2 semanas del mes de enero de 2010. La anterior diferencia en las condiciones de adquisición, genera desalineamientos en las señales bastante severos como se puede observar en la figura 15(a). Para corregir dicho desalineamiento, es necesario seleccionar una señal de referencia respecto a la cual serán alineadas todas las señales de cromatografía. La señal de referencia seleccionada para la base de datos de café fue una de las señales de la región 1 a la temperatura 3, pues mediante un algoritmo automático [Skov et al., 2006] se determinó que era la señal que más picos presenta, lo cual permite alinear de forma precisa los picos de las demás señales, aunque tengan una menor cantidad de éstos.

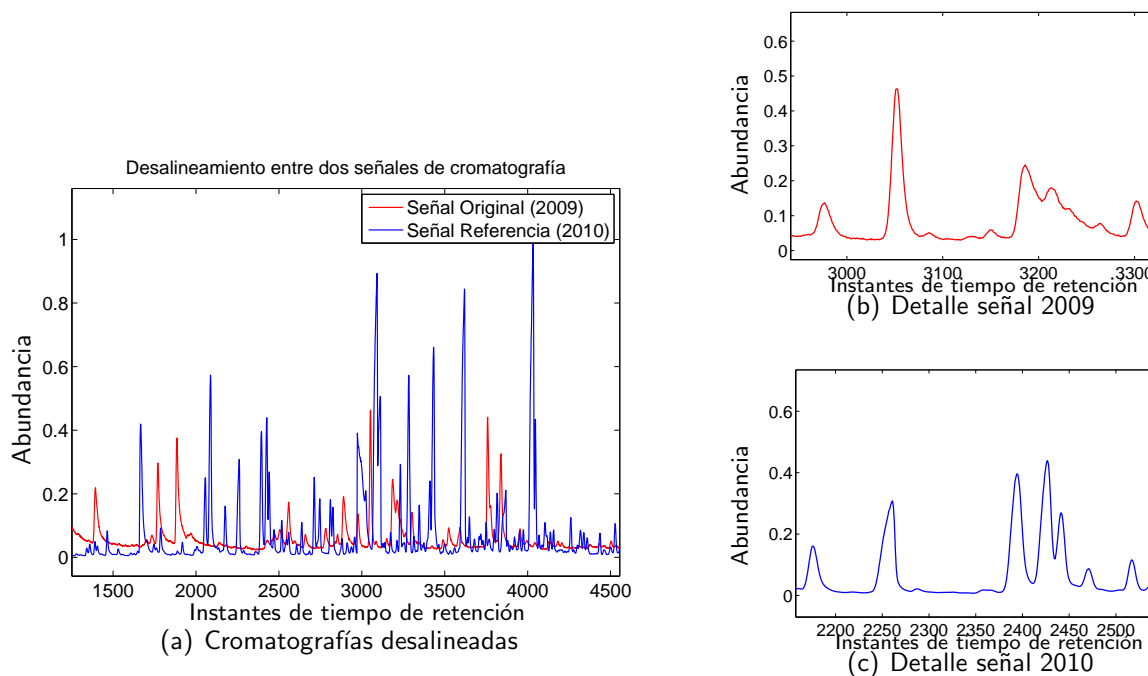


FIGURA 4.15. Señales típicas de cromatografía desalineadas

Nótese en las figuras 15(b) y 15(c) que picos correspondientes entre las dos señales fueron emitidos en instantes diferentes.

Inicialmente, fue necesario realizar un desplazamiento o corrimiento lineal de las señales del año 2009, tomando uno de los picos más representativos en todas ellas para alinearlas con respecto a dicho pico. Lo anterior se ejemplifica en la figura 4.16, donde se seleccionó como pico de referencia aquel emitido aproximadamente en el instante 2250.

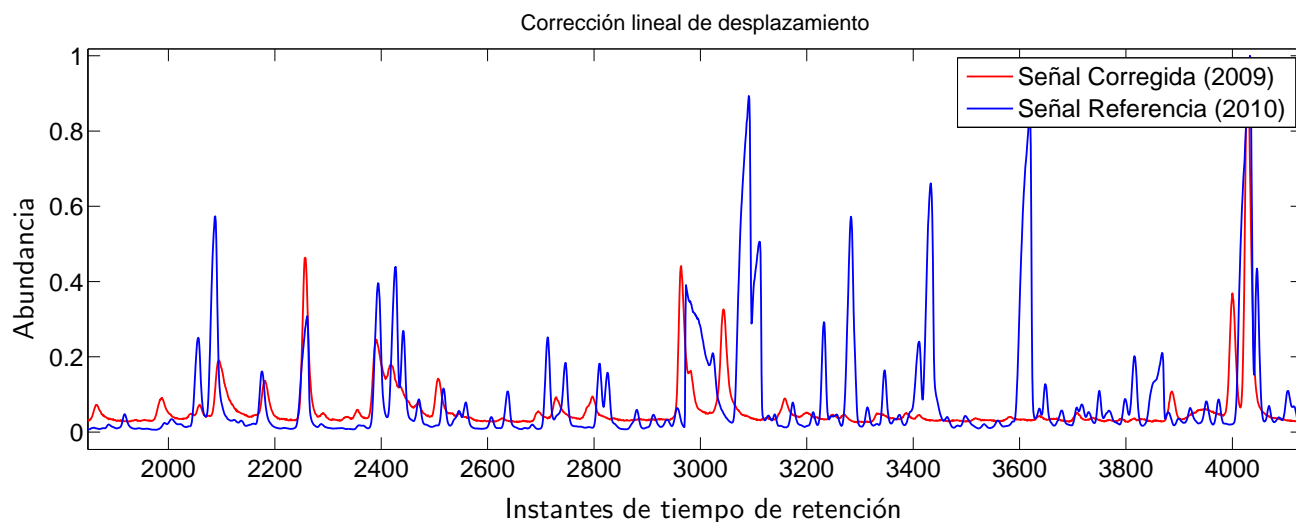


FIGURA 4.16. Resultado de la corrección lineal de desplazamiento.

En la literatura se ha reportado evidencia de la conveniencia de usar COW en lugar de DTW para señales de cromatografía [Tomasi et al., 2004]; en consecuencia, en este trabajo se ha optado por aplicar COW directamente. En las figuras 4.17 y 4.18 se ilustra el desempeño de COW bajo dos circunstancias particulares. En la figura 4.17 se logra exitosamente el alineamiento gracias a que originalmente los picos desalineados estaban relativamente cercanos. En contraste, en la figura 4.18 el algoritmo no logra alinear los picos puesto que los cambios de escala en porciones del eje horizontal son muy variables, tal como se observa en la primera mitad del eje horizontal de la figura.

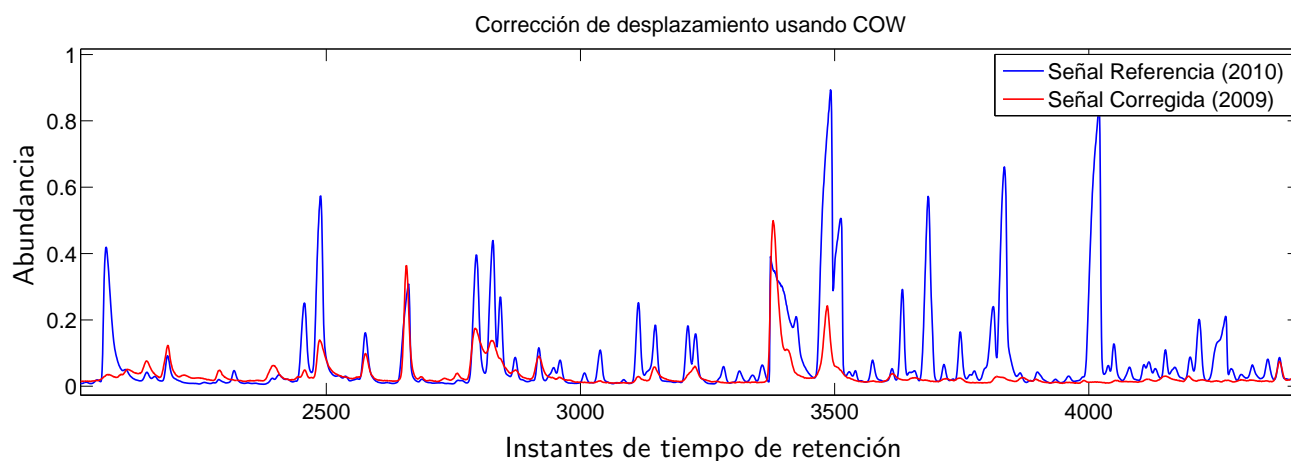


FIGURA 4.17. Porción de la señal exitosamente alineada usando el algoritmo COW

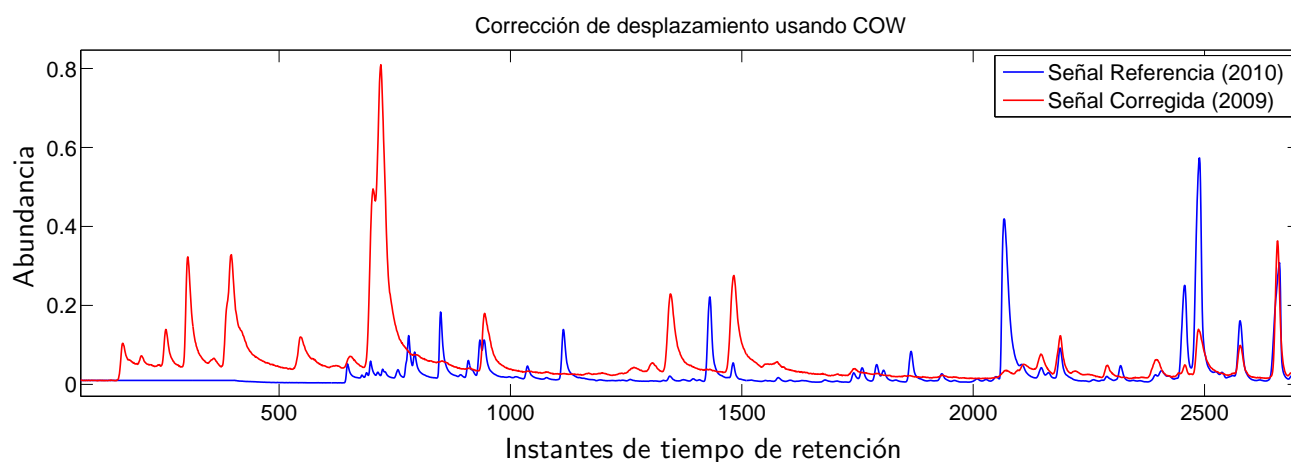


FIGURA 4.18. Porción de la señal con alineamiento fallido usando el algoritmo COW

Finalmente se aplica el algoritmo de ajuste de picos presentado en la sección 3.2.3. Mediante este algoritmo se logró alinear los cambios de escala bruscos en la parte inicial de las señales del 2009, tal que los picos coincidieran con las señales de 2010. El resultado del alineamiento se puede observar en la figura 4.19.



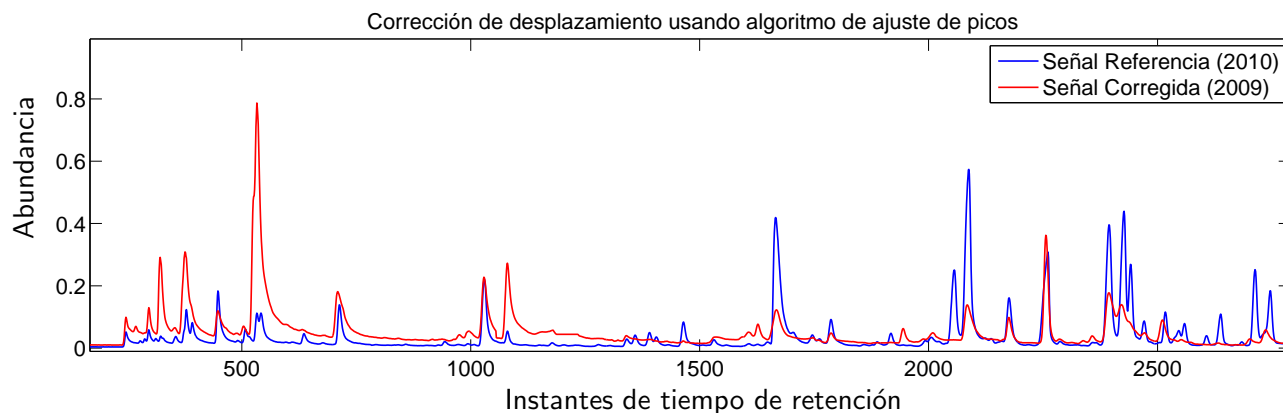


FIGURA 4.19. Porción de la señal exitosamente alineada usando ajuste de picos

Más de 1000 picos fueron observados en los cromatogramas de todas las regiones, pero solo 66 picos fueron encontrados en común a todas las muestras y usados como variables predictoras,  $X$ .

#### 4.6. Integración de datos sensoriales y señales de cromatografía

Inicialmente se aplica PCA sobre las 66 características de cromatografía de gases, que son 66 compuestos relevantes. Se calculan los componentes principales para todas las condiciones de región y temperatura, y se gráfica un Biplot de las muestras con las características para las dos primeras componentes principales.

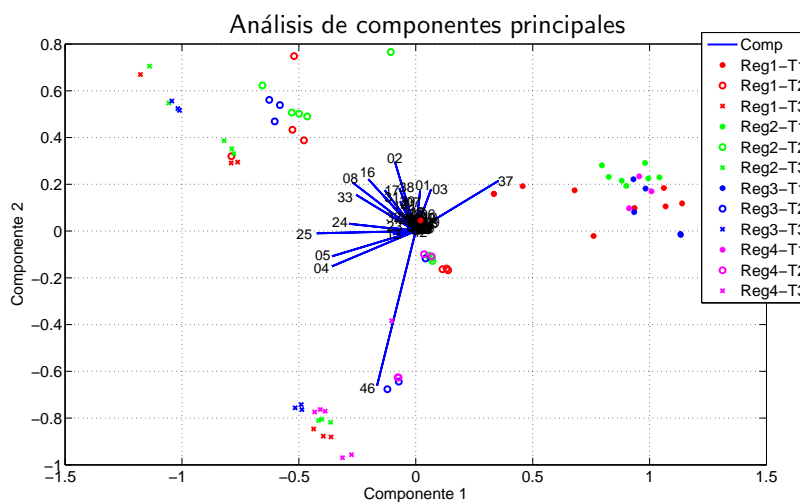


FIGURA 4.20. Biplot PCA para datos cromatográficos

En la gráfica 4.20 se observa la distribución de las muestras de las diferentes clases para las diferentes temperaturas proyectadas sobre las dos primeras componentes principales (PC).

También se grafican los pesos de las 66 características para las dos primeras PC. De esta forma se pueden encontrar relaciones entre las muestras y las características de acuerdo con las proyecciones obtenidas de PCA. Las características superpuestas en el origen de la gráfica indican su poca influencia en la mayor parte de variabilidad en los datos, mientras las características 46, 37, 4, 5, 25, 24, 33, 8, 16, 2, 1 y 3 muestran gran influencia en la variabilidad acumulada en las dos primeras PC. Los agrupamientos naturales de los datos corresponden a las temperaturas representadas en esta gráfica como asteriscos, cruces y círculos; que corresponden a cada una de las temperaturas. Mientras que las regiones, representadas por colores, se mezclan dificultándose la identificación automática de las regiones. Aplicando dos clasificadores simples, uno basado en la distribución de las clases, el clasificador discriminante lineal (LDA), y otro basado en la distancia, el clasificador del vecino más cercano (1-NN), se obtienen los resultados mostrados en la tabla 1:

Subconjunto de Características	Clases	Clasificador			
		LDA	LDA	1-NN	1-NN
		Verdaderos Posit. (%)	Falsos Positivos (%)	Verdaderos Posit. (%)	Falsos Positivos (%)
2 primeras PC	Región 1	29.13±1.13	34.07±1.00	59.13±1.33	10.00±0.79
PCA :95% de $\sigma^2$ acum		30.87±1.51	22.59±1.03	78.70±1.20	6.30±0.69
Dimensión Original		91.30±1.83	61.67±1.17	69.57±1.12	12.96±0.68
Selección Ida		69.57±2.34	10.37±1.08	73.91±1.12	10.74±0.25
2 primeras PC	Región 2	50.42±1.15	40.94±0.98	65.42±0.89	19.81±0.94
PCA :95% de $\sigma^2$ acum		44.17±1.88	23.96±1.05	70.42±2.19	20.00±0.50
Dimensión Original		22.08±2.41	7.17±1.19	67.08±1.58	20.00±0.75
Selección Ida		63.33±1.04	10.19±1.23	70.00±1.50	18.87±0.93
2 primeras PC	Región 3	0.00±0.00	0.17±0.17	65.88±1.92	9.83±0.63
PCA :95% de $\sigma^2$ acum		22.35±1.47	14.50±1.08	73.53±2.01	7.67±0.62
Dimensión Original		25.29±2.33	3.00±0.65	81.76±2.05	5.33±0.22
Selección Ida		74.12±1.30	11.33±0.60	51.18±1.26	8.50±0.39
2 primeras PC	Región 4	53.85±0.00	17.19±0.00	56.15±2.31	11.56±0.53
PCA :95% de $\sigma^2$ acum		37.69±2.13	26.56±0.04	53.85±2.29	6.09±0.89
Dimensión Original		34.62±2.36	4.69±0.09	60.77±2.42	3.59±0.62
Selección Ida		77.69±2.42	8.28±0.52	73.85±2.05	6.25±0.33
2 primeras PC	Temp. 1	96.15±0.00	0.00±0.00	96.15±0.00	0.00±0.00
PCA :95% de $\sigma^2$ acum		96.15±0.00	0.00±0.00	100±0.00	0.00±0.00
Dimensión Original		100±0.00	62.16±1.68	100±0.00	0.00±0.00
Selección Ida		100±0.00	2.75±0.32	100±0.00	0.00±0.00
2 primeras PC	Temp. 2	89.20±1.04	3.85±0.00	96.00±0.00	5.77±0.29
PCA :95% de $\sigma^2$ acum		96.00±0.00	1.92±0.00	96.00±0.00	5.19±0.29
Dimensión Original		43.60±2.34	1.92±0.81	96.00±0.00	6.92±0.43
Selección Ida		87.20±1.55	1.15±0.43	96.00±0.00	3.85±0.00
2 primeras PC	Temp. 3	96.15±0.00	5.29±0.51	92.31±0.57	1.96±0.00
PCA :95% de $\sigma^2$ acum		100±0.00	1.96±0.00	89.62±0.59	1.99±0.00
Dimensión Original		25.38±2.24	1.57±0.49	86.15±0.85	1.96±0.00
Selección Ida		97.69±0.85	3.53±0.64	92.31±0.00	1.96±0.00

a. Brockhoff – escalado multiplicativo

b. Brockhoff – escalado aditivo

En términos generales, la representación usando análisis de componentes principales no conduce a buenos resultados de clasificación para las regiones. En contraste, para las diferentes

temperaturas es una representación muy adecuada, esto se debe principalmente a que la variabilidad de los datos es causada en gran medida al efecto de las temperaturas sobre las características medidas. Para el clasificador discriminante lineal, aplicado a los datos proyectados en las dos primeras componentes principales (figura 4.20), los mejores desempeños de clasificación se obtienen para las clases región 2 y región 4. Los aciertos en la identificación de la clase región 2 se deben a que las muestras de la temperatura 3 de esta región están más cercanas al agrupamiento formado por las muestras de la temperatura 2, que al agrupamiento de las muestras de la temperatura 3 de las otras regiones. Similarmente, los aciertos en la identificación de muestras de la región 4 se deben a que las muestras de la temperatura 2 para esta región se ubican alrededor del origen, sin mostrar dependencia de algún agrupamiento, y cercanas a las muestras de las otras temperaturas de la misma clase. Las muestras de la clase región 3 son indetectables por este sistema de clasificación debido a que aparecen en la misma proporción en los tres agrupamientos, que a su vez están dispersos en todo el espacio de las dos componentes.

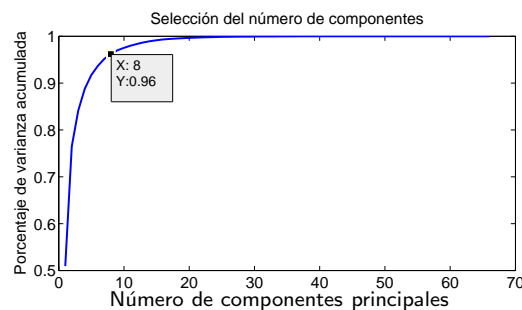


FIGURA 4.21. Varianza acumulada para los componentes principales

Adicionalmente, cuando se aplica la reducción de dimensión usando PCA, un criterio para escoger la cantidad de componentes se basa en la varianza acumulada, representada por los valores propios (figura 4.21). Según este criterio, con 8 componentes es suficiente para representar el 95% de la varianza total de los datos. El desempeño de clasificación, usando LDA y 8 factores PCA, no mejora considerablemente el desempeño de clasificación. Lo anterior puede deberse a que el criterio de reducción de dimensión de PCA basado en capturar la mayor cantidad de varianza enfatiza la variabilidad debida a las temperaturas y, por consiguiente, opaca la información representativa de las regiones.

Otro punto de referencia corresponde a aplicar el clasificador LDA sobre las características originales de los datos. Se observa que al usar esta representación gran cantidad de muestras pertenecientes a las demás clases son clasificadas como de la región 1. El desempeño de clasificación para la región 1 es muy alto 91.3%, pero también hay un porcentaje alto, 61.67%, de muestras erróneamente clasificadas como de esta región, que conduce a que el desempeño de clasificación de las otras regiones disminuya, a menos de 35%.

Otra representación propuesta consiste en realizar una selección de características hacia adelante basada en discriminación lineal, de la cual se obtiene una reducción de 66 características a 33. Esta representación conduce a un desempeño mejor y más balanceado en la discriminación de las regiones, con tasas de verdaderos positivos de alrededor de 70% para las diferentes regiones. De hecho, este tipo de representación es mucho más adecuada cuando se trata de discriminar las regiones, ya que se escoge un conjunto óptimo de características que permitan una discriminación mejor.

Por otro lado, si se observa los resultados de clasificación de las temperaturas usando la representación PCA, se aprecia que incluso con los dos primeros componentes los resultados de clasificación son iguales o mayores a 90% para las tres temperaturas consideradas, y usando 8 componentes que resumen más del 95% de la variabilidad de los datos se obtienen tasas de aciertos mayores a 95% para las tres temperaturas. Estos resultados ratifican el análisis de la proyección obtenida de las dos primeras PC (Figura 4.20) donde se evidencian tres agrupamientos bien separados correspondientes a las tres temperaturas. El desempeño decrece cuando se usan todas las características para discriminar estos tres grupos, lo cual puede deberse a que el 5% de variabilidad restante, representa ruido indeseable que de cierta forma opaca la separación obtenida mediante PCA. El último experimento con el clasificador LDA corresponde a la selección de características escogidas para discriminar entre las regiones, que también resulta beneficiosa en la identificación de las temperaturas, lo que implica que en estas características todavía hay un efecto muy importante de las temperaturas, que puede estar encubriendo la información discriminativa de las clases que nos interesa representar fidedignamente: las regiones.

Los resultados de clasificación, para las diferentes representaciones, usando el clasificador del vecino más cercano, presentan una mejoría considerable con respecto a los resultados obtenidos con el clasificador LDA. Esto se debe a que, aunque las regiones se distribuyan en igual proporción en los agrupamientos de las temperaturas, las muestras de una misma región se presentan en parejas o en grupos de tres, muy cercanas entre ellas, pero alejadas entre grupo y grupo. Estas pequeñas agrupaciones formadas en los datos se pueden deber a un tercer factor que no se mencionó inicialmente, el año en que fueron tomadas las muestras. Las representaciones que muestran mejores resultados, medidas de desempeño usando el clasificador 1-NN, es la de las 33 características seleccionadas junto con la representación original; Sin embargo, la selección de características logra obtener resultados más consistentes para todas las clases, siendo comparables con los resultados obtenidos del clasificador LDA con esta misma representación.

#### **4.6.1. Regresión por mínimos cuadrados parciales (PLS)**

Se aplica el análisis de regresión por mínimos cuadrados parciales con el objetivo de encontrar relaciones entre las características obtenidas de la cromatografía de gases y el análisis sensorial. Se obtienen tres representaciones, una diferente para cada catador. Este método proporciona parejas de componentes que consisten en dos factores, uno por cada tipo de datos, las varianzas acumuladas para cada factor de los modelos y cada uno de los catadores se presenta en la tabla 4.24. Para las tres representaciones los componentes que corresponden a los datos de cromatografía o variables predictivas, proporcionan representaciones casi

idénticas de los datos, mientras que las variables dependientes, los análisis sensoriales suministrados por los catadores, tienen representaciones muy diferentes, y esto demuestra la diferencia en la percepción de los atributos de cada uno de los catadores. El catador 1, por ejemplo, solo percibe 6 de los nueve atributos, puesto que en su representación de los atributos tres de ellos no tienen peso en las dos primeras componentes. Por otro lado, el catador 2 interpreta de forma muy diferente cada uno de los atributos y los usa para representar en menor o en mayor medida las regiones y temperaturas. El catador 3, por otro lado usa todos los atributos pero su percepción de algunos de éstos tiene una alta dependencia de otros, esto quiere decir que usa diferentes atributos para explicar una misma percepción en el aroma pero en diferentes proporciones.

Las relaciones encontradas usando regresión PLS se presentan en las tablas 4.22 - 4.23.

TABLA 4.22. Relaciones identificadas entre temperatura y compuestos usando PLS

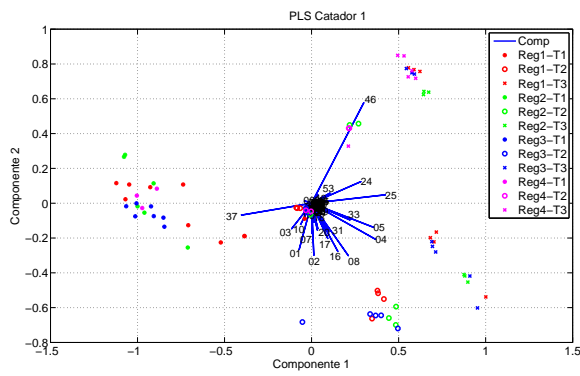
Temperatura	Compuesto	Aroma relacionado
Temp. 1	5-methyl 2-furancarboxaldehyde	caramelo
	4-vinylphenol 2-methoxy	picante y a clavos
Temp. 2	acetate 2-furanmethanol	frutal: banano
	ethyl pirazyne	maní, nueces, madera y chocolate
	2-methyl 3-pentanone	desconocido
	trimethyl pyrazine	papa, cacao, maní y asado
	ethanone 1-(1H-pyrrol-2-yl)	nueces, regaliz dulce, té dulce y tostado
Temp. 3	Pyrazine, 2-ethyl-6-methyl	tostado
	5-pyrazine 2-6 dimethyl	cocido o asado
	1H-pyrrole, 1-2(2-furanylmethyl)	vegetales, tierra, cereal, rábano y champiñones
	2-Butenoic acid 3-methyl	desconocido
	Furan 2-[(methyl thio) methyl]	café y tostado
	2,3 Butanedione	cremoso y mantequilla

TABLA 4.23. Relaciones identificadas entre temperatura y catadores usando PLS

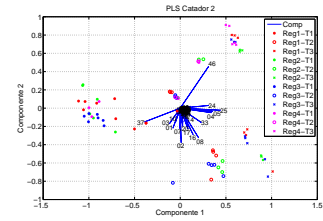
Temperatura	Catador	Atributo
Temp. 1	Catador 1	herbal, floral y madera
	Catador 2	floral, frutal, herbal, dulce
	Catador 3	herbal, madera y cereal
Temp. 2	Catador 1	madera y dulce
	Catador 2	cereal, madera y cítrico
	Catador 3	cereal y cítrico
Temp. 3	Catador 1	tostado
	Catador 2	tostado y cigarrillo
	Catador 3	tostado y dulce

TABLA 4.24. Varianza acumulada en las primeras componentes calculadas con PLS

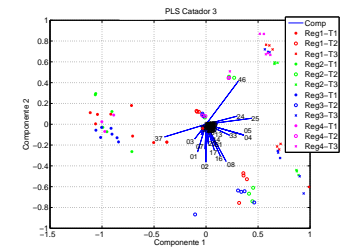
Componente	Catador 1		Catador 2		Catador 3	
	$\sigma^2$ acum de X (%)	$\sigma^2$ acum de Y (%)	$\sigma^2$ acum de X (%)	$\sigma^2$ acum de Y (%)	$\sigma^2$ acum de X (%)	$\sigma^2$ acum de Y (%)
1	52.16	72.91	50.27	46.18	50.15	62.80
2	77.12	76.25	76.12	49.76	75.70	66.25



(a) Modelo PLS Catador 1

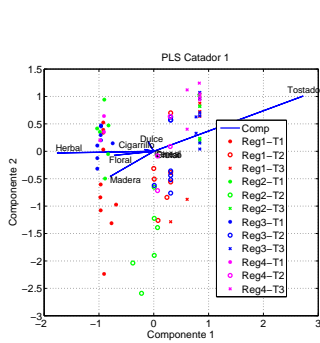


(b) Modelo PLS Catador 2

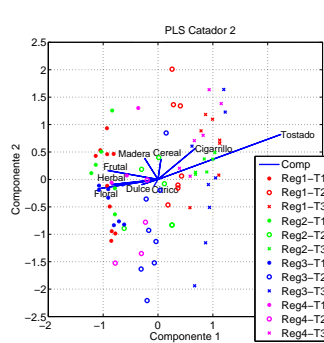


(c) Modelo PLS Catador 3

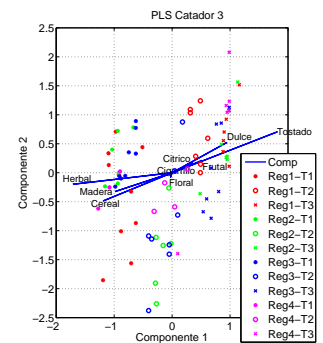
FIGURA 4.22. Datos transformados con el modelo PLS



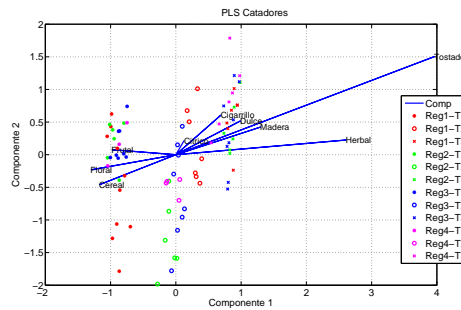
(a) Modelo PLS Catador 1



(b) Modelo PLS Catador 2



(c) Modelo PLS Catador 3



(d) Modelo PLS Catadores

FIGURA 4.23. Datos transformados con el modelo PLS matriz Y

#### 4.6.2. Correlación canónica (CCA)

Como resultado de la aplicación de la técnica de correlación canónica (sección 3.3.2) se obtuvieron las relaciones mostradas en las tablas 4.25, 4.26 y 4.27, obtenidas para cada uno de los tres catadores.

Los resultados de correlación canónica generan 9 funciones canónicas con sus respectivos coeficientes de correlación. Inicialmente, se observa a partir de los coeficientes la alta capacidad de las variables predictoras para explicar las variables criterio, puesto que las primeras 4 funciones canónicas presentan coeficientes de correlación de 1, lo cual es un indicio de la alta inter-correlación entre los datos de análisis sensorial y los compuestos químicos. Continuando con el proceso de análisis, se utilizaron varios criterios para decidir la cantidad de funciones canónicas que ayudan a interpretar la relación entre los dos conjuntos de variables. Con el primer criterio consiste en emplear la prueba de hipótesis basada en el lambda de Wilks la cual estableció que únicamente la primera variable canónica era significativa en la interpretación de la relación a un nivel de 0.05. El segundo criterio consiste en obtener un nivel de redundancia de la varianza compartida del conjunto de variables dependientes que puede ser explicada por la variable canónica independiente. Este índice de redundancia se calcula multiplicando la cantidad de varianza de las variables dependientes contenida en la variable canónica criterio y el coeficiente de correlación canónica. Los índices de redundancia para la primera y segunda función canónica son de 0.3193 y 0.0407 respectivamente. Según el análisis basado en estos criterios se decide analizar la primera función canónica obtenida.

Puesto que la relación canónica resultó estadísticamente significativa y las magnitudes de los coeficientes de correlación y del índice de redundancia son adecuados, la interpretación que comprende el examen de la función canónica determina la importancia relativa de cada una de las variables originales involucradas en las relaciones canónicas.

Tabla 4.25. Correlación canónica para el primer catador

ponderación	atributo	compuesto	ponderación	aroma
-1.66	herbal	Pyradine	17.02	Aroma fuerte a pescado, sabores a marisco, sabores ahumados, chocolate
-1.61	tostado	1-methyl 1H-Pyrrole	-18.99	Aroma a tostado
-0.61	dulce	Acetic acid	13.70	Vinagre o vino rancio
0.18	floral	Pyrazine, trimethyl	-11.07	Aroma con tonos a papa, cacao, maní y asado
0.00	madera	Phenol, 3-methyl-	-15.81	Aroma a medicina, a leña o a cuero
0.00	cigarrillo	Pyrazine, 2,6-dimethyl-	-9.12	Aroma con tonos a tostado
0.00	frutal	2-Furanmethanol	9.52	Aroma a horneado, tostado, nueces, almendras, ligero aroma a azúcar quemada
0.00	cereal	2-Butenoic acid, 3-methyl-	9.64	Aroma con tonos a café y a tostado
0.00	cítrico	2-Methoxy-4-vinylphenol	-8.23	Aroma con tonos picantes y a clavos
		2-Methoxy-4-vinylphenol	7.91	Aroma con tonos picantes y a clavos



Tabla 4.26. Correlación canónica para el segundo catador

ponderación	atributo	compuesto	ponderación	aroma
1.35	cigarrillo	2-Thiophenemethanol	35.63	No tiene aroma
0.84	frutal	1-methyl 1H-Pyrrole	33.87	Aroma a tostado
-0.53	tostado	Pyradine	-30.18	Aroma fuerte a pescado, sabores a marisco, sabores ahumados, chocolate
-1.25	floral	Acetic acid	23.76	Vinagre o vino rancio
0.87	frutal	Pyrazine, 2,6-dimethyl-	-22.54	Aroma con tonos a tostado
0.62	cigarrillo	1-Hydroxy-2-butanone	22.77	Notas a Madera, dulce, caramelo y nueces
-0.42	tostado	Furan, 2,2''-[oxybis(methylene)]bis	-21.42	No se encontró su aroma
0.24	cereal	2-Methoxy-4-vinylphenol	-20.50	Aroma con tonos picantes y a clavos
-0.18	dulce	Furan, 2,2''-methylenebis-	-20.92	No se reporta su aroma
0.18	cítrico	Phenol, 3-methyl-	16.92	Aroma a medicina, a leña o a cuero
-0.04	herbal	3(2H)-Furanone, dihydro-2-methyl-	15.28	conocido también como furanone de café. Aroma a dulce y a solventes con matices a nueces. Además su sabor es astringente con un matiz ligero a almendra cremoso.

Tabla 4.27. Correlación canónica para el tercer catador

ponderación	atributo	ponderación	compuesto	aroma
1.57	dulce	Furan, 2,2"-methylenebis-	-34.71	Aroma a quemado
1.26	cereal	2-Methoxy-4-vinylphenol	-28.08	Aroma con tonos picantes y a clavos
-0.79	tostado	Pyrazine, 2-ethyl-5-methyl-	-26.22	Conocido también como pyrazine de café. Aroma a grano de café tostado, dulce.
-0.53	madera	1H-Pyrrole, 1-(2-furanylmethyl)-	-23.51	No se reporta su aroma
-0.47	herbal	Pyrazine, trimethyl	22.45	Aroma con tonos a papa, cacao, maní y asado
-0.38	cítrico	2,5-Furandione, 3,4-dimethyl-	20.68	No se reporta su aroma
-0.30	floral	Phenol	-20.02	Empalagosamente dulce y alquitranado.
0.16	cigarrillo	2-Methoxy-4-vinylphenol	-20.50	Aroma con tonos picantes y a clavos.
-0.027	frutal	3-Pentanone, 2-methyl	-18.43	No se reporta su aroma
		2-Thiophenemethanol	17.21	No tiene aroma
		1-methyl 1H-Pyrrole	16.19	Aroma a tostado
		2-Cyclopenten-1-one, 3-ethyl-2-hydroxy-	15.59	Aroma a caramelo y coco, sabores dulces seductores.

Los siguientes compuestos son reconocidos como potencializadores del aroma del café, los cuales fueron identificados en las relaciones con los atributos sensoriales mediante el uso de PLS y CCA.

- Acetic acid [Ginz et al., 2000] [Schenker et al., 2002]
- 2-furanmethanol [Deibler and Delwiche, 2003]
- Pyridine [Andueza et al., 2003]
- 3-methyl phenol [Andueza et al., 2003]
- 2-thiophene methanol Phenol [Andueza et al., 2003]
- 2-Butenoic acid, 3-methyl- [Gonzalez-Rios et al., 2007]
- 2-Furanmethanol, acetate [Lopez-Galilea et al., 2008]
- pyrazine, 2,6-dimethyl- [Kim et al., 2000]
- phenol [Baggenstoss et al., 2008]
- furan, 2,2"-methylenebis- [Lopez-Galilea et al., 2008] [Sanz et al., 2001]
- 1-hydroxy 2- butanone [Gonzalez-Rios et al., 2007]
- 2-Methoxy-4-vinylphenol [Toci and Farah, 2008]

## 5. Discusión

Los resultados presentados en el capítulo anterior permiten sintetizar las siguientes observaciones generales:

- Las diferencias inherentes a los catadores fueron confirmadas mediante un modelo ANOVA de una vía (tabla 4.4).
- El modelo ANOVA de 3 vías para los datos originales muestra que el efecto del catador es el factor dominante para todos los atributos, excepto para el atributo tostado, seguido por el efecto de la temperatura (apéndice A).
- El efecto del catador es eliminado efectivamente mediante las transformaciones Procrustes y Brockhoff. (sección 4.3.3)
- La proyección sobre parejas de componentes principales permite una visualización aproximada de la distribución de las observaciones correspondientes a cada factor, a saber: temperatura, catador y región. (figuras 4.2–4.13).
- Las agrupaciones de catadores y de temperaturas son evidentes, mientras que no se pueden distinguir grupos de observaciones de cada región (figuras 4.2, 4.3 y 4.4).
- Los resultados de ANOVA de tres vías de las correcciones de diferencias individuales, cuyas variables fueron previamente combinadas con PCA, demuestran que se elimina el efecto del catador sobre los datos (sección 4.3.3).
- La estrategia de combinación de clasificadores ofrece la mejor alternativa en la identificación de regiones (sección 4.4.4).
- La transformación PCA permite visualizar los datos corregidos. En el caso de las temperaturas, se siguen observando agrupaciones; pero en el caso de las regiones, no se forman grupos, solo se puede resaltar que se eliminó el efecto de los catadores y particularmente, uno de ellos tiene una mayor dispersión en comparación con los otros dos (figuras 4.5–4.13).
- Los análisis PLS y CCA permitieron encontrar relaciones entre los atributos evaluados por el panel sensorial y los compuestos volátiles, que según la literatura, se consideran compuestos potencializadores del aroma del café secciones 4.6.1 y 4.6.2.



## 6. Conclusión

En este trabajo se presentó una metodología para el análisis de datos sensoriales e instrumentales. Específicamente, para identificar relaciones que indiquen la dependencia de los atributos sensoriales con respecto a los compuestos químicos potencializadores del aroma.

Con base en los resultados experimentales, una arquitectura recomendada para el sistema de identificación de relaciones entre datos sensoriales e instrumentales debería incluir las siguientes etapas: i) Adquisición de datos usando un protocolo basado en las normas técnicas colombianas, ver sección 4.2, ii) preprocesamiento de datos instrumentales empleando el algoritmo de ajuste de picos, seguido por alineamiento de correlación optimizado COW, ver secciones 3.2.3 y 3.2.2; asimismo, preprocesamiento de datos sensoriales mediante la corrección de la diferencia de evaluadores usando análisis de Procrustes, ver sección 3.1.8; finalmente, iii) Integración de datos sensoriales e instrumentales empleando regresión por mínimos cuadrados parciales para identificar las relaciones, ver sección 3.3.1.

La aplicación principal estuvo orientada a estudiar estas relaciones en muestras de café de origen; sin embargo, debido a inconvenientes logísticos por parte de la contraparte de este proyecto, fue necesario la consideración de ejemplos similares de otros productos alimenticios, en particular papa y queso. El mismo inconveniente mencionado hizo necesario enfatizar la etapa de preprocesamiento en lugar del objetivo inicialmente planteado de búsqueda de relaciones puesto que resultó imprescindible resolver inconsistencias y desacuerdos entre catadores. Se propuso la aplicación de transformaciones y correcciones para remover estas discrepancias, además de la creación de comités de clasificadores. Se encontró que esta última estrategia es la más conveniente para identificar diferencias entre productos. Además, el desempeño de combinación de clasificadores con regla de combinación de veto confirma que cada catador tiene habilidades superiores para identificar ciertas regiones particulares. Este hecho puede explicar también el pobre desempeño, en comparación con la combinación de clasificadores, obtenido tras las correcciones y transformaciones; puesto que la remoción del efecto del catador, al intentar homogenizar las evaluaciones de los catadores, tiende a deteriorar siempre las capacidades individuales.

Los experimentos usando PLS y CCA permitieron encontrar relaciones lineales entre los atributos sensoriales y compuestos potencializadores del aroma. Si bien se identificaron relaciones entre atributos y compuestos, se requiere de experimentos adicionales para lograr una predicción confiable de los atributos a partir de los compuestos químicos.

El desarrollo general del proyecto evidenció la gran importancia del seguimiento de un protocolo estandarizado de generación de datos sensoriales que no fue seguido en este trabajo

por la contraparte. Errores e inconsistencias en esta etapa son difícilmente superables en etapas posteriores mediante transformaciones o manipulaciones de corrección. Como de hecho en este caso, esta situación afectó no solo el cronograma del proyecto sino también la confiabilidad de las relaciones encontradas y la caracterización de las regiones.

Como trabajo futuro, usando los mismos conjuntos de datos de esta tesis, se propone aplicar técnicas de análisis categórico, dado que la escala que se empleó en la adquisición de datos sensoriales se basa en calificaciones que no indican realmente una cantidad sino la presencia o ausencia de un atributo. Teniendo en cuenta lo anterior, para conocer la relación entre atributos sensoriales y variables instrumentales, podrían usarse métodos de análisis multivariado de datos mixtos, tales como análisis factorial múltiple y clasificadores mixtos que emplean medidas de disimilitud apropiadas; es decir, que tienen en cuenta la naturaleza de los variables categóricas y cuantitativas simultáneamente.

## A. ANOVA para cada atributo

TABLA A.1. Comparación de corrección del atributo Floral con ANOVA

Factor	Suma de cuadrados	grados de libertad	Error cuadrático medio	F	valor p
<b>ANOVA de una vía para catadores</b>					
catador	162.31	2	81.15	203.12	0.0000
Error	106.68	267	0.39		
Total	269	269			
<b>ANOVA de tres vías: catador, temperatura y región</b>					
<b>Sin corrección</b>					
temp	51.88	2	25.94	135.19	0.0000
catador	204.37	2	102.18	532.54	0.0000
región	2.17	3	0.72	3.76	0.0115
temp*catador	26.64	4	6.66	34.71	0.0000
temp*región	3.98	6	0.66	3.46	0.0027
catador*región	3.96	6	0.66	3.44	0.0029
temp*catador*región	8.60	12	0.72	3.73	0.0000
Error	44.90	234	0.19		
Total	360.07	269			
<b>Procrustes</b>					
temp	133.23	2	66.61	350.55	0.0000
catador	0.12	2	0.06	0.32	0.7296
región	1.28	3	0.43	2.25	0.0834
temp*catador	0.85	4	0.21	1.12	0.3463
temp*región	4.55	6	0.76	3.99	0.0008
catador*región	3.03	6	0.50	2.66	0.0164
temp*catador*región	3.07	12	0.26	1.35	0.1939
Error	44.47	234	0.19		
Total	197.47	269			
<b>Modelo de evaluador</b>					
temp	107.38	2	53.69	15.37	0.0000
catador	0.18	2	0.09	0.03	0.9746
región	29.28	3	9.76	2.79	0.0411
temp*catador	13.85	4	3.46	0.99	0.4129
temp*región	36.82	6	6.14	1.76	0.1089
catador*región	49.11	6	8.18	2.34	0.0323
temp*catador*región	57.52	12	4.79	1.37	0.1800
Error	817.29	234	3.49		
Total	1113.16	269			
<b>Modelo de evaluador</b>					
temp	51.88	2	25.94	135.19	0.0000
catador	204.37	2	102.18	532.54	0.0000
región	2.17	3	0.72	3.76	0.0115
temp*catador	5.17	4	1.29	6.74	0.0000
temp*región	3.98	6	0.66	3.46	0.0027
catador*región	3.96	6	0.66	3.44	0.0029
temp*catador*región	8.60	12	0.72	3.73	0.0000
Error	44.90	234	0.19		
Total	336.42	269			

TABLA A.2. Comparación de corrección del atributo frutal usando ANOVA

Factor	Suma de cuadrados	grados de libertad	Error cuadrático medio	F	valor p
<b>ANOVA de una vía para catadores</b>					
catador	199.07	2	99.53	380.04	0.0000
Error	69.93	267	0.26		
Total	269	269			
<b>ANOVA de tres vías: catador, temperatura y región</b>					
<b>Sin corrección</b>					
temp	3.27	2	1.64	8.76	0.0002
catador	253.07	2	126.54	677.30	0.0000
región	1.71	3	0.57	3.05	0.0295
temp*catador	29.39	4	7.35	39.33	0.0000
temp*región	1.46	6	0.24	1.31	0.2548
catador*región	1.66	6	0.28	1.49	0.1840
temp*catador*región	5.15	12	0.43	2.30	0.0088
Error	43.72	234	0.19		
Total	354.97	269			
<b>Procrustes</b>					
temp	70.68	2	35.34	229.80	0.0000
catador	0.02	2	0.01	0.07	0.9335
región	2.05	3	0.68	4.45	0.0046
temp*catador	0.87	4	0.22	1.42	0.2284
temp*región	2.53	6	0.42	2.74	0.0135
catador*región	1.75	6	0.29	1.89	0.0831
temp*catador*región	3.84	12	0.32	2.08	0.0190
Error	35.98	234	0.15		
Total	124.63	269			
<b>Modelo de evaluador</b>					
temp	33.30	2	16.65	35.97	0.0000
catador	0.01	2	0.01	0.01	0.9864
región	0.40	3	0.13	0.29	0.8337
temp*catador	17.68	4	4.42	9.55	0.0000
temp*región	5.59	6	0.93	2.01	0.0646
catador*región	4.55	6	0.76	1.64	0.1374
temp*catador*región	6.44	12	0.54	1.16	0.3135
Error	108.30	234	0.46		
Total	186.33	269			
<b>Modelo de evaluador 2</b>					
temp	3.27	2	1.64	8.76	0.0002
catador	253.07	2	126.54	677.30	0.0000
región	1.71	3	0.57	3.05	0.0295
temp*catador	0.26	4	0.07	0.35	0.8435
temp*región	1.46	6	0.24	1.31	0.2548
catador*región	1.66	6	0.28	1.49	0.1840
temp*catador*región	5.15	12	0.43	2.30	0.0088
Error	43.72	234	0.19		
Total	320.12	269			



TABLA A.3. Comparación de corrección del atributo Tostado con ANOVA

Factor	Suma de cuadrados	grados de libertad	Error cuadrático medio	F	valor p
<b>ANOVA de una vía para catadores</b>					
catador	2.31	2	1.15	1.16	0.3162
Error	266.69	267	0.99		
Total	269	269			
<b>ANOVA de tres vías: catador, temperatura y región</b>					
<b>Sin corrección</b>					
temp	680.57	2	340.28	1076.03	0.0000
catador	8.14	2	4.07	12.87	0.0000
región	6.72	3	2.24	7.08	0.0001
temp*catador	19.54	4	4.89	15.45	0.0000
temp*región	5.19	6	0.87	2.74	0.0137
catador*región	7.65	6	1.28	4.03	0.0007
temp*catador*región	8.14	12	0.68	2.14	0.0152
Error	74.00	234	0.32		
Total	840.97	269			
<b>Procrustes</b>					
temp	545.45	2	272.72	1044.12	0.0000
catador	0.02	2	0.01	0.03	0.9714
región	8.57	3	2.86	10.94	0.0000
temp*catador	7.33	4	1.83	7.02	0.0000
temp*región	6.06	6	1.01	3.86	0.0011
catador*región	3.84	6	0.64	2.45	0.0257
temp*catador*región	6.82	12	0.57	2.17	0.0136
Error	61.12	234	0.26		
Total	664.71	269			
<b>Modelo de evaluador</b>					
temp	86.03	2	43.02	949.21	0.0000
catador	0.00	2	0.00	0.01	0.9877
región	1.11	3	0.37	8.15	0.0000
temp*catador	0.12	4	0.03	0.65	0.6282
temp*región	0.64	6	0.11	2.37	0.0308
catador*región	1.09	6	0.18	4.01	0.0008
temp*catador*región	0.97	12	0.08	1.78	0.0525
Error	10.60	234	0.05		
Total	104.52	269			
<b>Modelo de evaluador 2</b>					
temp	680.57	2	340.28	1076.03	0.0000
catador	8.14	2	4.07	12.87	0.0000
región	6.72	3	2.24	7.08	0.0001
temp*catador	0.95	4	0.24	0.75	0.5560
temp*región	5.19	6	0.87	2.74	0.0137
catador*región	7.65	6	1.28	4.03	0.0007
temp*catador*región	8.14	12	0.68	2.14	0.0152
Error	74.00	234	0.32		
Total	821.71	269			

TABLA A.4. Comparación de corrección del atributo Dulce con ANOVA

Factor	Suma de cuadrados	grados de libertad	Error cuadrático medio	F	valor p
<b>ANOVA de una vía para catadores</b>					
catador	164.52	2	82.26	210.22	0.0000
Error	104.47	267	0.39		
Total	269	269			
<b>ANOVA de tres vías: catador, temperatura y región</b>					
<b>Sin corrección</b>					
temp	4.76	2	2.38	8.46	0.0003
catador	182.08	2	91.04	323.43	0.0000
región	4.34	3	1.45	5.14	0.0019
temp*catador	40.46	4	10.12	35.94	0.0000
temp*región	3.17	6	0.53	1.88	0.0857
catador*región	7.63	6	1.27	4.52	0.0002
temp*catador*región	3.90	12	0.32	1.15	0.3177
Error	65.87	234	0.28		
Total	334.97	269			
<b>Procrustes</b>					
temp	4.73	2	2.37	11.23	0.0000
catador	0.03	2	0.02	0.08	0.9227
región	1.58	3	0.53	2.49	0.0608
temp*catador	0.92	4	0.23	1.10	0.3590
temp*región	1.74	6	0.29	1.37	0.2263
catador*región	2.52	6	0.42	1.99	0.0680
temp*catador*región	2.66	12	0.22	1.05	0.4027
Error	49.33	234	0.21		
Total	63.15	269			
<b>Modelo de evaluador</b>					
temp	110.58	2	55.29	14.86	0.0000
catador	6.13	2	3.06	0.82	0.4402
región	53.35	3	17.78	4.78	0.0030
temp*catador	8.85	4	2.21	0.59	0.6666
temp*región	57.78	6	9.63	2.59	0.0190
catador*región	209.00	6	34.83	9.36	0.0000
temp*catador*región	53.66	12	4.47	1.20	0.2825
Error	870.60	234	3.72		
Total	1398.32	269			
<b>Modelo de evaluador 2</b>					
temp	4.76	2	2.38	8.46	0.0003
catador	182.08	2	91.04	323.43	0.0000
región	4.34	3	1.45	5.14	0.0019
temp*catador	0.45	4	0.11	0.40	0.8059
temp*región	3.17	6	0.53	1.88	0.0857
catador*región	7.63	6	1.27	4.52	0.0002
temp*catador*región	3.90	12	0.32	1.15	0.3177
Error	65.87	234	0.28		
Total	296.19	269			

TABLA A.5. Comparación de corrección del atributo Cereal usando ANOVA

Factor	Suma de cuadrados	grados de libertad	Error cuadrático medio	F	valor p
<b>ANOVA de una vía para catadores</b>					
catador	125.35	2	62.67	116.51	0.0000
Error	143.64	267	0.53		
Total	269	269			
<b>ANOVA de tres vías: catador, temperatura y región</b>					
<b>Sin corrección</b>					
temp	18.09	2	9.04	35.06	0.0000
catador	127.44	2	63.72	247.07	0.0000
región	4.71	3	1.57	6.08	0.0005
temp*catador	42.17	4	10.54	40.88	0.0000
temp*región	3.39	6	0.57	2.19	0.0445
catador*región	12.71	6	2.12	8.21	0.0000
temp*catador*región	2.31	12	0.19	0.75	0.7058
Error	60.35	234	0.26		
Total	277.87	269			
<b>Procrustes</b>					
temp	5.32	2	2.66	10.82	0.0000
catador	0.07	2	0.04	0.14	0.8660
región	14.25	3	4.75	19.33	0.0000
temp*catador	5.12	4	1.28	5.20	0.0005
temp*región	6.38	6	1.06	4.33	0.0004
catador*región	6.55	6	1.09	4.44	0.0003
temp*catador*región	2.81	12	0.23	0.95	0.4956
Error	57.51	234	0.25		
Total	97.32	269			
<b>Modelo de evaluador</b>					
temp	1031.38	2	515.69	2.32	0.1004
catador	57.26	2	28.63	0.13	0.8791
región	7115.24	3	2371.75	10.68	0.0000
temp*catador	1916.90	4	479.23	2.16	0.0746
temp*región	1648.41	6	274.73	1.24	0.2882
catador*región	14150.41	6	2358.40	10.62	0.0000
temp*catador*región	3405.63	12	283.80	1.28	0.2326
Error	51977.87	234	222.13		
Total	80937.41	269			
<b>Modelo de evaluador</b>					
temp	18.09	2	9.04	35.06	0.0000
catador	127.44	2	63.72	247.07	0.0000
región	4.71	3	1.57	6.08	0.0005
temp*catador	1.39	4	0.35	1.35	0.2533
temp*región	3.39	6	0.57	2.19	0.0445
catador*región	12.71	6	2.12	8.21	0.0000
temp*catador*región	2.31	12	0.19	0.75	0.7058
Error	60.35	234	0.26		
Total	235.23	269			

TABLA A.6. Comparación de corrección del atributo Madera usando ANOVA

Factor	Suma de cuadrados	grados de libertad	Error cuadrático medio	F	valor p
<b>ANOVA de una vía para catadores</b>					
catador	91.91	2	45.95	69.29	0.0000
Error	177.09	267	0.66		
Total	269	269			
<b>ANOVA de tres vías: catador, temperatura y región</b>					
<b>Sin corrección</b>					
temp	69.21	2	34.60	129.89	0.0000
catador	87.70	2	43.85	164.58	0.0000
región	5.57	3	1.86	6.97	0.0002
temp*catador	18.97	4	4.74	17.80	0.0000
temp*región	7.07	6	1.18	4.43	0.0003
catador*región	4.52	6	0.75	2.83	0.0112
temp*catador*región	6.75	12	0.56	2.11	0.0170
Error	62.34	234	0.27		
Total	271.50	269			
<b>Procrustes</b>					
temp	6.85	2	3.43	21.35	0.0000
catador	0.03	2	0.01	0.09	0.9180
región	3.39	3	1.13	7.04	0.0002
temp*catador	1.36	4	0.34	2.12	0.0786
temp*región	2.10	6	0.35	2.18	0.0453
catador*región	2.21	6	0.37	2.29	0.0360
temp*catador*región	1.48	12	0.12	0.77	0.6824
Error	37.56	234	0.16		
Total	55.18	269			
<b>Modelo de evaluador</b>					
temp	86.76	2	43.38	33.91	0.0000
catador	0.12	2	0.06	0.05	0.9542
región	22.41	3	7.47	5.84	0.0007
temp*catador	4.22	4	1.05	0.82	0.5107
temp*región	10.72	6	1.79	1.40	0.2170
catador*región	26.48	6	4.41	3.45	0.0028
temp*catador*región	11.53	12	0.96	0.75	0.7006
Error	299.38	234	1.28		
Total	463.65	269			
<b>Modelo de evaluador 2</b>					
temp	69.21	2	34.60	129.89	0.0000
catador	87.70	2	43.85	164.58	0.0000
región	5.57	3	1.86	6.97	0.0002
temp*catador	3.15	4	0.79	2.96	0.0207
temp*región	7.07	6	1.18	4.43	0.0003
catador*región	4.52	6	0.75	2.83	0.0112
temp*catador*región	6.75	12	0.56	2.11	0.0170
Error	62.34	234	0.27		
Total	255.29	269			

TABLA A.7. Comparación de corrección del atributo Herbal usando ANOVA

Factor	Suma de cuadrados	grados de libertad	Error cuadrático medio	F	valor p
<b>ANOVA de una vía para catadores</b>					
catador	22.32	2	11.16	12.08	0.0000
Error	246.67	267	0.92		
Total	269	269			
<b>ANOVA de tres vías: catador, temperatura y región</b>					
<b>Sin corrección</b>					
temp	259.12	2	129.56	581.44	0.0000
catador	33.82	2	16.91	75.88	0.0000
región	1.80	3	0.60	2.69	0.0473
temp*catador	39.78	4	9.95	44.63	0.0000
temp*región	3.36	6	0.56	2.51	0.0225
catador*región	2.59	6	0.43	1.93	0.0762
temp*catador*región	12.92	12	1.08	4.83	0.0000
Error	52.14	234	0.22		
Total	418.39	269			
<b>Procrustes</b>					
temp	61.02	2	30.51	217.93	0.0000
catador	0.00	2	0.00	0.01	0.9919
región	0.78	3	0.26	1.85	0.1397
temp*catador	1.02	4	0.26	1.82	0.1251
temp*región	4.17	6	0.70	4.97	0.0001
catador*región	0.21	6	0.04	0.25	0.9577
temp*catador*región	5.82	12	0.49	3.47	0.0001
Error	32.76	234	0.14		
Total	107.17	269			
<b>Modelo de evaluador</b>					
temp	90.27	2	45.13	262.57	0.0000
catador	0.01	2	0.01	0.03	0.9692
región	0.62	3	0.21	1.20	0.3089
temp*catador	2.81	4	0.70	4.09	0.0032
temp*región	2.92	6	0.49	2.83	0.0112
catador*región	0.59	6	0.10	0.57	0.7549
temp*catador*región	8.58	12	0.71	4.16	0.0000
Error	40.22	234	0.17		
Total	149.46	269			
<b>Modelo de evaluador 2</b>					
temp	259.12	2	129.56	581.44	0.0000
catador	33.82	2	16.91	75.88	0.0000
región	1.80	3	0.60	2.69	0.0473
temp*catador	6.76	4	1.69	7.59	0.0000
temp*región	3.36	6	0.56	2.51	0.0225
catador*región	2.59	6	0.43	1.93	0.0762
temp*catador*región	12.92	12	1.08	4.83	0.0000
Error	52.14	234	0.22		
Total	72.23	269			

TABLA A.8. Comparación de corrección del atributo Cigarrillo usando ANOVA

Factor	Suma de cuadrados	grados de libertad	Error cuadrático medio	F	valor p
<b>ANOVA de una vía para catadores</b>					
catador	89.09	2	44.54	66.11	0.0000
Error	179.90	267	0.67		
Total	269	269			
<b>ANOVA de tres vías: catador, temperatura y región</b>					
<b>Sin corrección</b>					
temp	5.17	2	2.58	18.99	0.0000
catador	30.94	2	15.47	113.72	0.0000
región	1.18	3	0.39	2.89	0.0360
temp*catador	18.96	4	4.74	34.85	0.0000
temp*región	0.41	6	0.07	0.50	0.8087
catador*región	0.83	6	0.14	1.02	0.4152
temp*catador*región	1.68	12	0.14	1.03	0.4219
Error	31.83	234	0.14		
Total	93.47	269			
<b>Procrustes</b>					
temp	57.30	2	28.65	194.33	0.0000
catador	0.00	2	0.00	0.00	0.9990
región	1.70	3	0.57	3.85	0.0103
temp*catador	1.34	4	0.33	2.27	0.0624
temp*región	1.04	6	0.17	1.17	0.3226
catador*región	0.62	6	0.10	0.70	0.6514
temp*catador*región	0.71	12	0.06	0.40	0.9614
Error	34.50	234	0.15		
Total	101.99	269			
<b>Modelo de evaluador</b>					
temp	77.71	2	38.85	5.19	0.0062
catador	0.81	2	0.40	0.05	0.9475
región	33.83	3	11.28	1.51	0.2133
temp*catador	7.58	4	1.89	0.25	0.9075
temp*región	81.22	6	13.54	1.81	0.0979
catador*región	28.96	6	4.83	0.65	0.6938
temp*catador*región	42.06	12	3.51	0.47	0.9316
Error	1750.18	234	7.48		
Total	1972.69	269			
<b>Modelo de evaluador 2</b>					
temp	5.17	2	2.58	18.99	0.0000
catador	30.94	2	15.47	113.72	0.0000
región	1.18	3	0.39	2.89	0.0360
temp*catador	0.12	4	0.03	0.21	0.9311
temp*región	0.41	6	0.07	0.50	0.8087
catador*región	0.83	6	0.14	1.02	0.4152
temp*catador*región	1.68	12	0.14	1.03	0.4219
Error	31.83	234	0.14		
Total	193.81	269			

TABLA A.9. Comparación de corrección del atributo Cítrico usando ANOVA

Factor	Suma de cuadrados	grados de libertad	Error cuadrático medio	F	valor p
<b>ANOVA de una vía para catadores</b>					
catador	173.85	2	86.92	243.95	0.0000
Error	95.14	267	0.35		
Total	269	269			
<b>ANOVA de tres vías: catador, temperatura y región</b>					
<b>Sin corrección</b>					
temp	2.11	2	1.06	5.48	0.0047
catador	129.99	2	64.99	337.27	0.0000
región	0.73	3	0.24	1.26	0.2887
temp*catador	18.39	4	4.60	23.85	0.0000
temp*región	2.69	6	0.45	2.33	0.0335
catador*región	0.68	6	0.11	0.59	0.7399
temp*catador*región	3.94	12	0.33	1.70	0.0666
Error	45.09	234	0.19		
Total	209.10	269			
<b>Procrustes</b>					
temp	13.57	2	6.79	40.96	0.0000
catador	0.00	2	0.00	0.01	0.9933
región	1.08	3	0.36	2.16	0.0931
temp*catador	4.42	4	1.11	6.68	0.0000
temp*región	4.55	6	0.76	4.57	0.0002
catador*región	0.94	6	0.16	0.94	0.4638
temp*catador*región	3.04	12	0.25	1.53	0.1144
Error	38.77	234	0.17		
Total	66.89	269			
<b>Modelo de evaluador</b>					
temp	54.38	2	27.19	20.00	0.0000
catador	0.01	2	0.00	0.00	0.9970
región	2.31	3	0.77	0.57	0.6375
temp*catador	43.31	4	10.83	7.97	0.0000
temp*región	15.94	6	2.66	1.95	0.0731
catador*región	8.22	6	1.37	1.01	0.4211
temp*catador*región	38.53	12	3.21	2.36	0.0069
Error	318.07	234	1.36		
Total	480.37	269			
<b>Modelo de evaluador</b>					
temp	2.11	2	1.06	5.48	0.0047
catador	129.99	2	64.99	337.27	0.0000
región	0.73	3	0.24	1.26	0.2887
temp*catador	1.84	4	0.46	2.38	0.0524
temp*región	2.69	6	0.45	2.33	0.0335
catador*región	0.68	6	0.11	0.59	0.7399
temp*catador*región	3.94	12	0.33	1.70	0.0666
Error	45.09	234	0.19		
Total	193.81	269			

## Bibliografía

- [Adhikari et al., 2003] Adhikari, K., Heymann, H., and Huff, H. E. (2003). Textural characteristics of lowfat, fullfat and smoked cheeses: sensory and instrumental approaches. *Food Quality and Preference*, 14(3):211 – 218.
- [Aishima, 2004] Aishima, T. (2004). Correlating sensory attributes to gas chromatography-mass spectrometry profiles and e-nose responses using partial least squares regression analysis. *Journal of Chromatography A*, 1054(1-2):39–46.
- [Altman and Bland, 2005] Altman, D. G. and Bland, J. M. (2005). Standard deviations and standard errors. *BMJ*, 331(7521):903.
- [Andueza et al., 2003] Andueza, S., Maeztu, L., Pascual, L., Ibáñez, C., de Peña, M. P., and Cid, C. (2003). Influence of extraction temperature on the final quality of espresso coffee. *Journal of the Science of Food and Agriculture*, 83(3):240–248.
- [Arvanitoyannis et al., 2007] Arvanitoyannis, I. S., Mavromatis, A., Rodiatis, A., and Goulas, C. (2007). Physicochemical and sensory analysis of dry bean landraces in conjunction with multivariate analysis: an exploratory approach. *International Journal of Food Science and Technology*, 42(7):819–826.
- [Aznar et al., 2003] Aznar, M., Lopez, R., Cacho, J., and Ferreira, V. (2003). Prediction of aged red wine aroma properties from aroma chemical composition. partial least squares regression models. *Journal of Agricultural and Food Chemistry*, 51(9):2700–2707.
- [Baggenstoss et al., 2008] Baggenstoss, J., Poisson, L., Kaegi, R., Perren, R., and Escher, F. (2008). Coffee roasting and aroma formation: Application of different time-temperature conditions. *Journal of Agricultural and Food Chemistry*, 56(14):5836–5846.
- [Bramer, 2007] Bramer, M. (2007). *Principles of Data Mining*. Springer.
- [Bro et al., 2008] Bro, R., Qannari, E. M., Kiers, H. A. L., Næs, T., and Frøst, M. B. (2008). Multi-way models for sensory profiling data. *Journal of Chemometrics*, 22(1):36–45.
- [Brockhoff, 2003] Brockhoff, P. B. (2003). Statistical testing of individual differences in sensory profiling. *Food Quality and Preference*, 14(5-6):425 – 434. The Sixth Sense - 6th Sensometrics Meeting.
- [Brockhoff, 1998] Brockhoff, P. M. (1998). Assessor modelling. *Food Quality and Preference*, 9(3):87 – 89. Sensometric Workshop.
- [Brockhoff and Skovgaard, 1994] Brockhoff, P. M. and Skovgaard, I. M. (1994). Modelling individual differences between assessors in sensory evaluations. *Food Quality and Preference*, 5(3):215 – 224.
- [Chung et al., 2003] Chung, S.-J., Heymann, H., and Grün, I. U. (2003). Application of gpa and pls-r in correlating sensory and chemical data sets. *Food Quality and Preference*, 14(5-6):485–495. The Sixth Sense - 6th Sensometrics Meeting.
- [Cozzolino et al., 2006] Cozzolino, D., Smyth, H. E., Lattey, K. A., Cynkar, W., Janik, L., Damberg, R. G., Francis, I. L., and Gishen, M. (2006). Combining mass spectrometry



- based electronic nose, visible-near infrared spectroscopy and chemometrics to assess the sensory properties of australian riesling wines. *Analytica Chimica Acta*, 563(1-2):319 – 324.
- [Deibler and Delwiche, 2003] Deibler, K. D. and Delwiche, J. (2003). *Handbook of flavor characterization: sensory analysis, chemistry, and physiology*. CRC Press.
- [Dijksterhuis, 1994] Dijksterhuis, G. (1994). Procrustes analysis in studying sensory-instrumental relations. *Food Quality and Preference*, 5(1-2):115 – 120.
- [Dijksterhuis, 1996] Dijksterhuis, G. (1996). Procrustes analysis in sensory research. In Næs, T. and Risvik, E., editors, *Multivariate analysis of data in sensory science*, volume 16 of *Data Handling in Science and Technology*, pages 185–219. Elsevier.
- [Du et al., 2010] Du, X., Kurnianta, A., McDaniel, M., Finn, C., and Qian, M. (2010). Flavour profiling of [']marion' and thornless blackberries by instrumental and sensory analysis. *Food Chemistry*, 121(4):1080 – 1088.
- [Durante et al., 2006] Durante, C., Cocchi, M., Grandi, M., Marchetti, A., and Bro, R. (2006). Application of n-pls to gas chromatographic and sensory data of traditional balsamic vinegars of modena. *Chemometrics and Intelligent Laboratory Systems*, 83(1):54 – 65.
- [François et al., 2008] François, I. M., Wins, H., Buysens, S., Godts, C., Pee, E. V., Nicolai, B., and Proft, M. D. (2008). Predicting sensory attributes of different chicory hybrids using physico-chemical measurements and visible/near infrared spectroscopy. *Postharvest Biology and Technology*, 49(3):366 – 373.
- [Ginz et al., 2000] Ginz, M., Balzer, H. H., Bradbury, A. G. W., and Maier, H. G. (2000). Formation of aliphatic acids by carbohydrate degradation during roasting of coffee. *European Food Research and Technology*, 211(6):404–410.
- [Gonzalez-Rios et al., 2007] Gonzalez-Rios, O., Suarez-Quiroz, M. L., Boulanger, R., Barel, M., Guyot, B., Guiraud, J.-P., and Schorr-Galindo, S. (2007). Impact of “ecological” post-harvest processing on coffee aroma: li. roasted coffee. *Journal of Food Composition and Analysis*, 20(3-4):297 – 307.
- [Granitto et al., 2008] Granitto, P., Biasioli, F., Endrizzi, I., and Gasperi, F. (2008). Discriminant models based on sensory evaluations: Single assessors versus panel average. *Food Quality and Preference*, 19(6):589 – 595.
- [Hair et al., 1998] Hair, J. F., Tatham, R. L., Anderson, R. E., and Black, W. (1998). *Multivariate Data Analysis*. Prentice Hall.
- [Höskuldsson, 988] Höskuldsson, A. (988). PLS regression methods. *Journal of Chemometrics*, 2(3):211–228.
- [ICONTEC, 1990] ICONTEC (1990). NTC 2680: Industrias alimentarias. Análisis sensorial. Ensayo de comparación por pares.
- [ICONTEC, 1996] ICONTEC (1996). NTC 3925: Análisis sensorial. Metodología. Guía general.
- [ICONTEC, 1997a] ICONTEC (1997a). NTC 2679: Análisis sensorial. Metodología. Ensayo “A” y “no A”.
- [ICONTEC, 1997b] ICONTEC (1997b). NTC 4129: Análisis sensorial. Guía general para la selección, entrenamiento y seguimiento de evaluadores. Parte 1. Evaluadores seleccionados.
- [ICONTEC, 2000] ICONTEC (2000). NTC 4883: Análisis sensorial. Café. Metodología para análisis sensorial cuantitativo descriptivo del café.

- [ICONTEC, 2002] ICONTEC (2002). NTC 2758: Café. Análisis sensorial. Vocabulario.
- [ICONTEC, 2006] ICONTEC (2006). NTC 2681: Análisis sensorial. Metodología. Prueba triangular.
- [Johnson et al., 2003] Johnson, K. J., Wright, B. W., Jarman, K. H., and Synovec, R. E. (2003). High-speed peak matching algorithm for retention time alignment of gas chromatographic data for chemometric analysis. *Journal of Chromatography A*, 996(1-2):141 – 155.
- [Johnson and Wichern, 2002] Johnson, R. A. and Wichern, D. W. (2002). *Applied Multivariate Statistical Analysis*. Prentice Hall.
- [Jørgensen and Næs, 2008] Jørgensen, K. and Næs, T. (2008). The use of ls-pls for improved understanding, monitoring and prediction of cheese processing. *Chemometrics and Intelligent Laboratory Systems*, 93(1):11 – 19.
- [Kallithraka et al., 2001] Kallithraka, S., Arvanitoyannis, I. S., Kefalas, P., El-Zajouli, A., Soufleros, E., and Psarra, E. (2001). Instrumental and sensory analysis of greek wines; implementation of principal component analysis (pca) for classification according to geographical origin. *Food Chemistry*, 73(4):501 – 514.
- [Kermit and Lengard, 2005] Kermit, M. and Lengard, V. (2005). Assessing the performance of a sensory panel-panellist monitoring and tracking. *Journal of Chemometrics*, 19(3):154–161.
- [Kim et al., 2000] Kim, S., Yoon, H., and Rhee, J. (2000). The effects of roasting temperatures on the formation of headspace volatile compounds in perilla seed oil. *Journal of the American Oil Chemists' Society*, 77(4):451–456.
- [Kutner et al., 2005] Kutner, M. H., Nachtsheim, C. J., Neter, J., and Li, W. (2005). *Applied Linear Statistical Models*. McGraw Hill, 5th edition.
- [Lassoued et al., 2008] Lassoued, N., Delarue, J., Launay, B., and Michon, C. (2008). Baked product texture: Correlations between instrumental and sensory characterization using flash profile. *Journal of Cereal Science*, 48(1):133 – 143.
- [Lindinger et al., 2008] Lindinger, C., Labbe, D., Pollien, P., Rytz, A., Juillerat, M. A., Yeretzyan, C., and Blank, I. (2008). When machine tastes coffee: Instrumental approach to predict the sensory profile of espresso coffee. *Analytical Chemistry*, 80(5):1574–1581.
- [Lopez-Galilea et al., 2008] Lopez-Galilea, I., de Peña, M., and Cid, C. (2008). Application of multivariate analysis to investigate potential antioxidants in conventional and torrefacto roasted coffee. *European Food Research and Technology*, 227(1):141–149.
- [Luciano and Næs, 2009] Luciano, G. and Næs, T. (2009). Interpreting sensory data by combining principal component analysis and analysis of variance. *Food Quality and Preference*, 20(3):167 – 175.
- [Meilgaard et al., 2007] Meilgaard, M. C., Civille, G. V., and Carr, B. T. (2007). *Sensory evaluation techniques*. CRC Press.
- [Montgomery and Runger, 2002] Montgomery, D. C. and Runger, G. C. (2002). *Applied Statistics and Probability for Engineers*. John Wiley & Sons.
- [Moskowitz, 1996] Moskowitz, H. R. (1996). Sensory analysis procedures and viewpoints: intellectual history, current debates, future outlooks. *Journal of Sensory Studies*, 8(3):241–256.

- [Næs and Kowalski, 1989] Næs, T. and Kowalski, B. (1989). Predicting sensory profiles from external instrumental measurements. *Food Quality and Preference*, 1(4-5):135 – 147.
- [Nielsen et al., 1998] Nielsen, N.-P. V., Carstensen, J. M., and Smedsgaard, J. (1998). Aligning of single and multiple wavelength chromatographic profiles for chemometric data analysis using correlation optimised warping. *Journal of Chromatography A*, 805(1-2):17 – 35.
- [Parolari et al., 1992] Parolari, G., Virgili, R., Bolzoni, L., Careri, M., and Mangia, A. (1992). Analysis of sensory and instrumental data on apricot purées with pattern recognition techniques. *Analytica Chimica Acta*, 259(2):257 – 265.
- [Peña, 2002] Peña, D. (2002). *Análisis de datos multivariantes*. Mc Graw Hill.
- [Pierce et al., 2005] Pierce, K. M., Hope, J. L., Johnson, K. J., Wright, B. W., and Synovec, R. E. (2005). Classification of gasoline data obtained by gas chromatography using a piecewise alignment algorithm combined with feature selection and principal component analysis. *Journal of Chromatography A*, 1096(1-2):101 – 110. Chemical Separations and Chemometrics.
- [P.W.Scott, 1998] P.W.Scott, R. (1998). *Introduction to analytical gas chromatography*, volume 76. Marcel Dekker.
- [Ribeiro et al., 2009] Ribeiro, J. S., Augusto, F., Salva, T. J. G., Thomaziello, R. A., and Ferreira, M. M. C. (2009). Prediction of sensory properties of brazilian arabica roasted coffees by headspace solid phase microextraction-gas chromatography and partial least squares. *Analytica Chimica Acta*, 634(2):172–179.
- [Romano et al., 2008] Romano, R., Brockhoff, P. B., Hersleth, M., Tomic, O., and Næs, T. (2008). Correcting for different use of the scale and the need for further analysis of individual differences in sensory analysis. *Food Quality and Preference*, 19(2):197 – 209.
- [Sanz et al., 2001] Sanz, C., Ansorena, D., Bello, J., and Cid, C. (2001). Optimizing headspace temperature and time sampling for identification of volatile compounds in ground roasted arabica coffee. *Journal of Agricultural and Food Chemistry*, 49(3):1364–1369.
- [Scacco et al., 2007] Scacco, A., Mazzaglia, A., Todaro, A., and Lanza, C. M. (2007). Sensory and Physicochemical Characterization of Cerasuolo di Vittoria Red Wine. *American Journal of Enology and Viticulture*, 58(1):112–116.
- [Schenker et al., 2002] Schenker, S., Heinemann, C., Huber, M., Pompizzi, R., Perren, R., and Escher, R. (2002). Impact of roasting conditions on the formation of aroma compounds in coffee beans. *Journal of Food Science*, 67(1):60–66.
- [Schulbach et al., 2004] Schulbach, K., Rouseff, R., and Sims, C. (2004). Relating descriptive sensory analysis to gas chromatography/olfactometry ratings of fresh strawberries using partial least squares regression. *Journal of Food Science*, 69(7):273–277.
- [Skov et al., 2006] Skov, T., van den Berg, F., Tomasi, G., and Bro, R. (2006). Automated alignment of chromatographic data. *Journal of Chemometrics*, 20(11):484–497.
- [Thybo and Martens, 1998] Thybo, A. K. and Martens, M. (1998). Development of a sensory texture profile of cooked potatoes by multivariate data analysis. *Journal of Texture Studies*, 29(4):453–468.
- [Toci and Farah, 2008] Toci, A. T. and Farah, A. (2008). Volatile compounds as potential defective coffee beans' markers. *Food Chemistry*, 108(3):1133 – 1141.

- [Tomasi et al., 2004] Tomasi, G., van den Berg, F., and Andersson, C. (2004). Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. *Journal of Chemometrics*, 18(5):231 – 241.
- [Verzera et al., 2008] Verzera, A., Ziino, M., Scacco, A., Lanza, C., Mazzaglia, A., Romeo, V., and Condurso, C. (2008). Volatile compound and sensory analysis for the characterization of an italian white wine from “inzolia” grapes. *Food Analytical Methods*, 1(2):144–151.
- [Williams et al., 1988] Williams, A. A., Rogerst, C. A., and Collins, A. J. (1988). Relating chemical/physical and sensory data in food acceptance studies. *Food Quality and Preference*, 1(1):25 – 31.
- [Wold, 1995] Wold, S. (1995). Chemometrics; what do we mean with it, and what do we want from it? *Chemometrics and Intelligent Laboratory Systems*, 30(1):109–115.