Revista Colombiana de Matemáticas
Vol. XIX (1985), págs. 263-276

WEAK APPROXIMATION OF MINIMAL NORM SOLUTIONS OF FIRST KIND EQUATIONS BY TIKHONOV'S METHOD

by

Julio E. GUACANEME

ABSTRACT. Tikhonov's regularization method is considered to find conditions that guarantee orders of weak convergence of approximate solutions of linear ill-posed problems to the true solution. We establish orders of convergence by requiring smoothness conditions on the functional and the true solution, and we establish a converse result to the main theorem.

RESUMEN. Se establecen órdenes de convergencia débil para las soluciones aproximadas obtenidas por el método de regularización de Tikhonov en el caso de problemas lineales "ill-posed" (es decir, aquellos para los cuales las soluciones exactas pueden depender discontinuamente de los parámetros). Para ello se exigen condiciones de suavidad tanto al funcional como a la solución exacta. Esto se hace para la versión clásica infinito-dimensional del método de Tikhonov y también para la versión con elementos finitos. Además, se obtiene un converso al teorema principal, en el cual la suavidad resulta del orden de convergencia.

1. Introducción. In this article we shall be concerned with the integral equation of the first kind

$$Kx(s) = \int_0^1 k(s,t)x(t)dt = g(s) \qquad (1)$$

where $k(s,t) \in L^2([0,1]\times[0,1])$ and $g(s) \in L^2[0,1]$. It is well known that K is a compact linear operator from the Hilbert space $H_1 = L^2[0,1]$ into the Hilbert space $H_2 = L_2[0,1]$. Solving (1) then means finding an x satisfying (1), given $g \in H_2$.

This problem can have solutions x (we do not assume uniqueness; by solution we mean "minimal norm solution") which depend discontinuously on the "data" g, i.e., this is not a well-posed problem (see e.g. [6]). This lack of continuity can have serious numerical consequences since the data g is usually the result of measurements and hence is only imprecisely known.

Regarding the discontinuous dependence upon data, instead of solving (1) we solve a new equation close to (1) which is well posed. This approach is called "regularization". In particular Tikhonov [6] suggests the minimizer x_α of the functional

$$F_\alpha(w,g) = \|Kw-g\|^2 + \alpha\|w\|^2 \qquad (2)$$

as a regularized solution of (1); we use $\|\cdot\|$ to indicate the norm in each of the space H_1 and H_2 and $<\cdot,\cdot>$ to denote the corresponding inner product. The minimizer x_α of this functional is the solution of the equation

$$(\alpha I + \tilde{K})x_\alpha = K^* g,$$

where K^* denotes the adjoint of K and $\tilde{K} = K^*K$, see e.g. [4].

When the approximation x_α is defined, we would like to estimate how far it is from the minimum norm solution x. This can be done in the strong sense by looking at the norm of the difference $\|x-x_\alpha\|$ between the approximate and the true solutions, or in the weak sense by considering the functional $<x-x_\alpha,z>$, where $z \in H_1$.

In many applications all we want to know about solutions is the value of some functional $<x,z>$ (see e.g. [2]). In such cases we will be concerned with $<x,z>$ and $<x_\alpha,z>$ rather than x and x_α respectively. That is, our interest is in weak rather than strong approximation. In this paper we will derive estimates for $<x-x_\alpha,z>$ under various assumptions on x and z for both the classical infinite dimensional version of Tikhonov regularization and for a finite element version. We consider both the cases of exact data and inaccurate data.

2. Infinite Dimensional Tikhonov Approximation.

For fixed $z \in H_1$ we consider whether

$$<x-x_\alpha,z> \to 0 \quad \text{as } \alpha \to 0,$$

i.e., the weak convergence of the Tikhonov approximations to the true solution.

Sufficient conditions for convergence in the weak topology have already been studied by Tikhonov [7], and for more general methods of regularization, by H.W. Engl [1].

Our goal is to establish orders of convergence by imposing conditions on x and the functional z. Since x as well as x_α are members of $\overline{R(K^*)}$, the closure of the range of K^*, and every z in H has an orthogonal decomposition $z = z_1+z_2 \in \overline{R(K^*)} + R(K^*)^\perp$, we have

$$\langle x-x_\alpha, z\rangle = \langle x-x_\alpha, z_1\rangle .$$

We therefore may restrict our attention to funcionals $z \in \overline{R(K^*)}$; however to obtain convergence rates we need to impose stronger conditions on z.

THEOREM 2.1. *If* (a) $x \in R(\tilde{K})$ *or* (b) $z \in R(\tilde{K})$ *or* (c) x *and* $z \in R(K^*)$, *then* $\langle x-x_\alpha, z\rangle = O(\alpha)$.

Proof. (a) If $x \in R(\tilde{K})$, then $|\langle x-x_\alpha, z\rangle| \leq \|x-x_\alpha\|\|z\| = O(\alpha)$ by [4, Corollary 3.1.1]. (b) For this case let $z = Ku$, then

$$|\langle x-x_\alpha, z\rangle| = |\langle\tilde{K}(x-x_\alpha), u\rangle| = |\langle[I-\tilde{K}(\alpha I+\tilde{K})^{-1}]K^*g, u\rangle|$$

$$= \alpha|\langle(\alpha I+\tilde{K})^{-1}\tilde{K}x, u\rangle| \leq \alpha\|x\|\|u\|,$$

since $\|(\alpha I+\tilde{K})^{-1}\tilde{K}\| \leq 1$. (c) Let $z = K^*v$, $x = K^*w$, then, setting $\hat{K} := KK^*$, we have

$$|\langle x-x_\alpha, z\rangle| = |\langle K(x-x_\alpha), v\rangle = |\langle\hat{K}w-\hat{K}(\alpha I+K)^{-1}\hat{K}w, v\rangle|$$

$$= \alpha|\langle\hat{K}(\alpha I+\hat{K})^{-1}w, v\rangle| \leq \alpha\|w\|\|v\|,$$

which completes the proof. ▲

The best order of weak convergence is $O(\alpha)$ as the following theorem establishes.

THEOREM 2.2. *If* $\langle x-x_\alpha, z\rangle = O(\alpha)$ *for every* $z \in H_1$, *then* $x = 0$.

Proof. Since $<x-x_\alpha, z> = 0(\alpha)$, we have $<(x,x_\alpha)/\alpha, z> \to 0$ as $\alpha \to 0$ $\forall z \in H_1$. In particular the sequence $\{(x-x_\alpha)/\alpha\}$ is weakly convergent and hence bounded, i.e., $\|x-x_\alpha\| = 0(\alpha)$. By [4, Theorem 3.2.2] we have $x \in R(K^*K)$, say $x = K^*Kv$. Now using the fact that $x-x_\alpha = \alpha(\alpha I+\tilde{K})^{-1}\tilde{K}v$ (clearly we may take $v \in N(\tilde{K})^\perp = N(K)^\perp$), we then have

$$0 = \lim_{\alpha\to 0} <(x-x_\alpha)/\alpha, z> = \lim_{\alpha\to 0} (\alpha I+\tilde{K})^{-1}\tilde{K}v, z> = <K^+Kv, z>$$

for every $z \in H_1$, so $K^+Kv = 0$, where K^+ is the Moore-Penrose inverse of K (see e.g. [4]). But

$$K^+Kv = P_{N(K)^\perp}v = v$$

and hence $v = 0$, i.e., $x = \tilde{K}v = 0$. ▲

For the converse of Theorem 2.1 we will have, using the same notation as above, the following

THEOREM 2.3. *If* $<x-x_\alpha, z> = 0(\alpha)$ *for all* $z \in H_1$, *then* $x \in R(K^*K)$.

Proof. Define $T_\alpha z = <(x-x_\alpha)/\alpha, z>$. This family of linear functionals on H_1 has the properties required by the uniform boundedness theorem, that is,

$$\|T_\alpha z\| = |<\frac{x-x_\alpha}{\alpha}, z>| \leq M(z) \text{ for all } \alpha > 0.$$

M(z) is a bound depending of the element z, and thus there exists a bound M independent of α and independent of z such that $\|T_\alpha\| \leq M$, i.e.,

$$\|T_\alpha z\| = |<\frac{x-x_\alpha}{\alpha}, z>| \leq M\|z\|$$

for all $\alpha > 0$, and for all $z \in H_1$. In particular, for $z = (x-\alpha)/\alpha$,

$$\left\|T\frac{x-x_\alpha}{\alpha}\right\| = \left\|\frac{x-x_\alpha}{\alpha}\right\|^2 \leqslant M\left\|\frac{x-x_\alpha}{\alpha}\right\|,$$

and hence $\|x-x_\alpha\| = O(\alpha)$. By [4 Theorem 3.2.2] we conclude that $x \in R(K^*K)$. ▲

Now that we have analyzed the error-free case we turn our attention to the more realistic case of inexact data g^δ, with a prescribed error bound $\delta : \|g-g^\delta\| \leqslant \delta$.

The Tikhonov regularized solution x_α^δ is the minimizer of

$$F_\alpha(z,g^\delta) = \|Kz-g^\delta\|^2 + \alpha\|z\|^2,$$

or, equivalently, $x_\alpha^\delta = (\alpha I+\tilde{K})^{-1}K^*g^\delta$.

LEMMA 2.4. *If* $z \in R(K^*)$ *then* $\langle x_\alpha - x_\alpha^\delta, z\rangle = O(\delta)$ *for any* $\alpha > 0$.

Proof. Let $z = K^*u$, then

$$\langle x_\alpha - x_\alpha^\delta, z\rangle = \langle K(x_\alpha - x_\alpha^\delta), u\rangle = \langle \hat{K}(\alpha I+\hat{K})^{-1}(g-g^\delta), u\rangle$$

and hence

$$|\langle x_\alpha - x_\alpha^\delta, z\rangle| \leqslant \|g-g^\delta\|\|u\| \leqslant \delta\|u\|,$$

since $\|\hat{K}(\alpha I+\hat{K})^{-1}\| \leqslant 1$. ▲

THEOREM 2.5. *Let* $\alpha = O(\delta)$. *If* (a) $z \in R(\tilde{K})$, *or* (b) z *and* $x \in R(K^*)$, *then* $\langle x-x_\alpha^\delta, z\rangle = O(\delta)$.

Proof. (a) By Theorem 2.1 and the preceding lemma

$$|\langle x-x_\alpha^\delta, z\rangle| \leq |\langle x-x_\alpha, z\rangle| + |\langle x_\alpha - x_\alpha^\delta, z\rangle| = O(\alpha) + O(\delta) = O(\delta)$$

(b) Let $z = K^*w$ and $x = K^*v$, then

$$\langle x-x_\alpha, z\rangle = \langle K(x-x_\alpha), w\rangle = \alpha\langle(\alpha I+\hat{K})^{-1}\hat{K}v, w\rangle$$

and hence

$$|\langle x-x_\alpha, z\rangle| \leq \alpha\|v\|\|w\| = O(\alpha).$$

Using the Lemma again we find that $\langle x-x_\alpha^\delta, z\rangle = O(\alpha) + O(\delta)$. ▲

In the final theorem of this section, we make no assumption on the true solution x.

THEOREM 2.6. *If* $z \in R(K^*)$, *then* $\langle x-x_\alpha, z\rangle = O(\sqrt{\alpha})$.

Proof. Suppose $z = K^*u$, then

$$\langle x-x_\alpha, z\rangle = \langle K(x-x_\alpha), u\rangle.$$

But

$$\|K(x-x_\alpha)\|^2 = \langle K(x-x_\alpha), K(x-x_\alpha)\rangle = \langle\tilde{K}(x-x_\alpha), x-x_\alpha\rangle.$$

But $\tilde{K}(x-x_\alpha) = K^*g - \tilde{K}(\tilde{K}+\alpha I)^{-1}K^*g = \alpha\tilde{K}(\tilde{K}+\ I)^{-1}K^*g = \alpha\tilde{K}x_\alpha$. Therefore,

$$\|K(x-x_\alpha)\|^2 = \alpha\langle\tilde{K}x_\alpha, x-x_\alpha\rangle = O(\alpha),$$

since $x_\alpha \to x$ (see |4|). Therefore,

$$\langle x-x_\alpha, z\rangle = O(\|K(x-x_\alpha)\|) = O(\sqrt{\alpha}),$$

completing the proof. ▲

Combining this with Lemma 2.4 we obtain:

COROLLARY 2.7. *If* $z \in R(K^*)$ *and* $\alpha = O(\delta^2)$, *then* $<x-x_\alpha^\delta, z> = O(\delta)$. ▲

3. Finite Element Approximations. Groetsch and Guacaneme [3] have proved weak convergence of certain finite element Tikhonov approximation to the minimal norm solution of (1). These approximations are formed by using a sequence of finite dimensional subspaces V_m that increase and are eventually dense in H_1, i.e.,

$$V_1 \subseteq V_2 \subseteq \ldots \text{ and } \overline{\cup_m V_m} = H_1 .$$

The finite element Tikhonov approximations $x_{m,\alpha}$ and $x_{m,\alpha}^\delta$ are the minimizers of $F_\alpha(\cdot;g)$ and $F_\alpha(\cdot;g^\delta)$ (see [2]), respectivley, over the finite dimensional space V_m, or equivalently $x_{m,\alpha}, x_{m,\alpha}^\delta \in V_m$ and

$$<Kx_{m,\alpha}-g, Ky> + \alpha<x_{m,\alpha}, y> = 0$$

$$<Kx_{m,\alpha}-g^\delta, Ky> + <\alpha x_{m,\alpha}^\delta, y> = 0$$

respectively, for all $y \in V_m$. These conditions are in turn equivalent to

$$x_{m,\alpha} = (K_m^* K_m + \alpha I)^{-1} K_m^* g$$

and

$$x_{m,\alpha}^\delta = (K_m^* K_m + \alpha I)^{-1} K_m^* g^\delta$$

where K_m is the restriction of K to V_m. The number

$$\gamma_m = \|K(I-P_m)\| = \|(I-P_m)K^*\|$$

where P_m is the orthogonal projector of H_1 onto V_m, plays a prominent role; it tell us how well the spaces V_m support the operator K. Note that $\gamma_m \to 0$ as $m \to \infty$ (see e.g. [4]).

To study the approximations $x_{m,\alpha}$ and $x^{\delta}_{m,\alpha}$ we assume the regularization parameter is a function of m, say $\alpha = \alpha(m) \to 0$ as $m \to \infty$.

Before going into the order of convergence we define the inner product $[u,v] = \langle Ku,Kv\rangle + \alpha\langle u,v\rangle$ and the norm $|u|^2 = [u,u] = \|Ku\|^2 + \alpha\|u\|^2$ in the Hilbert space H_1. Under this inner product $x_{m,\alpha}$ is orthogonal projection of x_α onto V_m, (see [5]), i.e.

$$[x_\alpha - x_{m,\alpha}, v] = \langle \tilde{K}(x_\alpha - x_{m,\alpha}), v\rangle + \alpha\langle x_\alpha - x_{m,\alpha}, v\rangle = 0$$

for all $v \in V_m$.

We now give a weak order of convergence result for $x_{m,\alpha}$. For ease of notation below we will replace α_m by α and γ_m by γ, respectively.

THEOREM 3.1. *Assume that* $\gamma = O(\alpha^{\frac{1}{2}})$. (a) *If* x *and* $z \in R(K^*)$ *then* $\langle x - x_{m,\alpha}, z\rangle = O(\alpha)$. (b) *If* $z \in R(K^*)$, *then* $\langle x - x_{m,\alpha}, z\rangle = O(\sqrt{\alpha})$.

Proof. (a) Let $z = K^*w$ and let $x = K^*v$. Then,

$$\langle x - x_{m,\alpha}, z\rangle = \langle x - x_\alpha, z\rangle + \langle x_\alpha - x_{m,\alpha}, z\rangle$$

$$= \langle K(x - x_\alpha), w\rangle + \langle K(x_\alpha - x_{m,\alpha}), w\rangle.$$

Now,

$$K(x - x_\alpha) = \hat{K}v - \hat{K}(\alpha I + \hat{K})^{-1}\hat{K}v = \alpha(\alpha I + \hat{K})^{-1}\hat{K}v;$$

therefore

$$|\langle K(x-x_\alpha), w\rangle| \leq \alpha\|v\|\|w\| = 0(\alpha).$$

Also

$$\|K(x_\alpha - x_{m,\alpha})\|^2 \leq \|K(x_\alpha - x_{m,\alpha})\|^2 + \alpha\|x_\alpha - x_{m,\alpha}\|^2$$

and remembering that $x_{m,\alpha}$ is the projection of x_α onto V_m under the inner product $[w,v] = \langle Kw,Kv\rangle + \alpha\langle w,v\rangle$, we have

$$\begin{aligned}
\|K(x_\alpha - x_{m,\alpha})\|^2 + \alpha\|x_\alpha - x_{m,\alpha}\|^2 &\leq \|K(x_\alpha - P_m x_\alpha)\|^2 + \alpha\|x_\alpha - P_m x_\alpha\|^2 \\
&\leq \|K(I-P_m)^2 x_\alpha\|^2 + \alpha\|(I-P_m)x_\alpha\|^2 \\
&\leq (\gamma^2+\alpha)\|(I-P_m)x_\alpha\|^2 \\
&= (\gamma^2+\alpha)\|(I-P_m)K^*(\alpha I+\hat{K})^{-1}\hat{K}v\|^2 \\
&\leq (\gamma^2+\alpha)\gamma^2\|v\|^2.
\end{aligned}$$

Therefore

$$\begin{aligned}
|\langle x_\alpha - x_{m,\alpha}, z\rangle| &\leq \|K(x_\alpha - x_{m,\alpha})\|\|w\| \\
&\leq (\gamma^2+\alpha)^{\frac{1}{2}}\gamma\|v\|\|w\| = 0(\alpha).
\end{aligned}$$

(b) Let $z = Ku$, using the same decomposition as in part (a) we have

$$\langle x-x_{m,\alpha}, z\rangle = \langle x-x_\alpha, z\rangle + \langle x_\alpha - x_{m,\alpha}, z\rangle$$

and $\langle x-x_\alpha, z\rangle = 0(\sqrt{\alpha})$, by Theorem 2.6. Now $\langle x_\alpha - x_{m,\alpha}, z\rangle = \langle K(x_\alpha - x_m,\alpha), u\rangle$, and

$$\|K(x_\alpha - x_{m,\alpha})\|^2 \leq \|K(x_\alpha - x_{m,\alpha})\|^2 + \alpha\|x_\alpha - x_{m,\alpha}\|^2.$$

Using the characterization of $x_{m,\alpha}$ as the projection of x_α

onto V_m, we have as above

$$\|K(x_\alpha - x_{m,\alpha})\|^2 \leq (\gamma^2+\alpha)\|(I-P_m)x_\alpha\|^2 = O(\alpha)\|(I-P_m)x_\alpha\|^2 = O(\alpha),$$

thus

$$<x_\alpha - x_{m,\alpha}, z> = O(\sqrt{\alpha}).$$

We therefore find that

$$<x - x_{m,\alpha}, z> = (\sqrt{\alpha})$$

for case (b). ▲

We see that under suitable conditions the finite element approximations attain the order $O(\alpha)$ of weak convergence, which we now show does not allow improvement.

THEOREM 3.2. *If* $<x-x_{m,\alpha}, z> = O(\alpha)$ *for all* $z \in H_1$, *then* $x = 0$.

Proof. In particular for $z = \tilde{K}(u+w)$ with $u \in V_N$ and $w \in V_N^\perp$, we have

$$<x-x_{m,\alpha}, z> = O(\alpha).$$

By the definition of $x_{m,\alpha}$ we have

$$<\tilde{K}(x-x_{m,\alpha}), u> + \alpha<x_{m,\alpha}, u> = 0 \text{ for } m > N.$$

Therefore

$$<x-x_{m,\alpha}, \tilde{K}(u+w)> = \tilde{K}(x-x_{m,\alpha}), u+w>$$

$$= -\alpha<x_{m,\alpha}, u> + <\tilde{K}(x-x_{m,\alpha}), w>$$

and hence

$$0 = \lim_m \frac{<x-x_{m,\alpha}, \tilde{K}(u+w)>}{\alpha} = -\lim_m <x_{m,\alpha}, u> + \lim_m <\frac{\tilde{K}(x-x_{m,\alpha})}{\alpha}, w>$$

$$= -\lim_m <x_{m,\alpha}, u>,$$

for $u \in V_N$. To see this note that by hypothesis $\frac{x-x_{m,\alpha}}{\alpha}$ converges weakly to zero and K is a compact operator, therefore $\tilde{K}(\frac{x-x_{m,\alpha}}{\alpha})$ converges to zero. Using the fact that $x_{m,\alpha} \xrightarrow{w} x$, we then have $0 = <x,u>$ for all $u \in V_N$. Since $\bigcup_{n=1}^{\infty} V_n$ is dense in H_1 we have $x = 0$. ▲

Finally, we consider the weak convergence of the approximation $x^{\delta}_{m,\alpha}$ obtained using imprecise data g^{δ}, where $\|g-g^{\delta}\| < \delta$.

LEMMA 3.3. $\|K(x_{m,\alpha} - x^{\delta}_{m,\alpha})\| < \delta$.

Proof. Since

$$K(x_{m,\alpha}-x^{\delta}_{m,\alpha}) = K_m(\alpha I+\tilde{K}_m)^{-1}K^*_m(g-g^{\delta})$$

$$= \hat{K}_m(\alpha I+\hat{K}_m)^{-1}(g-g^{\delta})$$

and

$$\|\hat{K}_m(\alpha I+\hat{K}_m)^{-1}\| \leq 1$$

we get that

$$\|K(x_{m,\alpha}-x^{\delta}_{m,\alpha})\| \leq \|g-g^{\delta}\| \leq \delta. \ ▲$$

We now show that under appropiate conditions an optimal orde of weak convergence obtains.

THEOREM 3.4. *Assume that* $\gamma = O(\alpha^{½})$. *If either* (a) x *and* $z \in R(K^*)$ *and* $\alpha = O(\delta)$ *or* (b) $z \in R(K^*)$ *and*

$\alpha = O(\delta^2)$, *then* $\langle x - x_{m,\alpha}, z\rangle = O(\delta)$.

Proof. By Theorem 3.1, in both cases (a) and (b):

$$\langle x - x_{m,\alpha}, z\rangle = O(\delta).$$

Now set $z = K^*w$, then by Lemma 3.3 we have

$$|\langle x_{m,\alpha} - x^{\delta}_{m,\alpha}, z\rangle| = |\langle K(x_{m,\alpha} - x^{\delta}_{m,\alpha}), w\rangle| \leqslant \delta\|w\|,$$

i.e., $\langle x_{m,\alpha} - x_{m,\alpha}, z\rangle = O(\delta)$, and hence $\langle x - x^{\delta}_{m,\alpha}, z\rangle = O(\delta)$, completing the proof. ▲

REFERENCES

[1] Engl, H.W., *Necessary and sufficient conditions for convergence of regularization methods for solving linear operator equations of the first kind.* Numer. Funct. Anl. & Optimiz. 3 (1981), 201-222.

[2] Golberg, M.A., *A method of adjoints for solving ill-posed problems*. Appl. Math. & Comp., 5 (1979), 123-130.

[3] Groetsch, C.W. and Guacaneme, J., *On regularized Ritz approximations for Fredholm equations of the first kind.* Rocky Mountain Journal of Mathematics, 15 (1985), 33-39.

[4] Groetsch, C.W., *The theory of Tikhonov regularization for Fredholm equations of the first kind.* Pitman, London, 1984.

[5] Groetsch, C.W., King, J.T. and Murio D., *Asymptotic analysis of a finite element method for Fredholm equations of the first kind.* In *Treatment of integral equations by numerical methods*, (C.T. H. Baker and G.F. Miller, Eds.) Academic Press, London, 1982.

[6] Tikhonov, A.N. and Arsenin, V.Y., *Solutions of ill-posed problems*, Wiley, New York, 1977.

[7] Tikhonov, A.N., *Regularization of incorrectly posed problems*. Soviet Math. Doklady 4 (1963), 1624-1627.

Department of Mathematics
University of Puerto Rico
Río Piedras, Puerto Rico, 00931

(Recibido en octubre de 1984; versión revisada en mayo de 1985).