A NEWTON-GRADIENT METHOD FOR NON-LINEAR PROBLEMS IN HILBERT SPACE

By

Duane A. Meeter

Technical Report Number 7

NASA Grant Number NGR-10-004-029

August 24, 1966

Department of Statistics
Florida State University
Tallahassee, Florida

A NEWTON-GRADIENT METHOD FOR NON-LINEAR PROBLEMS IN HILBERT SPACE.

by

Duane A. Meeter

1. Introduction. In [5], Marquardt developed and extended an iterative method which had been proposed by others (see [4]) for non-linear least squares problems. Marquardt's paper demonstrates that the method is an interpolation between Newton's method (actually the Taylor Series or Gauss method for non-linear least squares) and the gradient or steepest descent method. The method produces a correction vector to the current iterate whose length and orientation is controlled by an adjustable parameter λ. Marquardt produced an algorithm for choosing λ at each iteration. The method has worked well on a variety of problems, e.g. [6]. It is generally more stable than the Taylor Series method and faster than the steepest descent method.

2. The Method in Hilbert Space. This paper extends the method to (real) Hilbert space. It turns out that each of the three theorems in Marquardt [5] has its counterpart in Hilbert space; however all of the proofs are new.

Let X and Y be real Hilbert spaces, F a non-linear operator $F: X \to Y$. Then the least squares problem may be stated as, for fixed $y \in Y$, minimize over X the functional

$$Q(x) = \| y-F(x) \|^2 = \langle y-F(x), y-F(x) \rangle \tag{2.1}$$

Let x_o be the first approximation to an x that minimizes Q. Let P be the Frechet derivative of F, with increment h; $P = P(x)h$, and let P^* denote the adjoint of P. Then the operator $A: X \to X$, where $A = P^*(x_o)P(x_o)$, is linear, self-adjoint, and positive. We assume that A is strictly positive and bounded, i.e., there exist constants β_1, β_2 such that $\beta_1 > \beta_2 > 0$, and

$$\beta_2 \langle x,x \rangle \leq \langle Ax,x \rangle \leq \beta_1 \langle x,x \rangle \tag{2.2}$$

for all $x \in X$.

The gradient of Q at x_o, increment h, is $-2\langle y-F(x_o), P(x_o)h \rangle$. Alternatively, apart from the constant -2, we can express the gradient at x_o as the functional

$$g = P^*(x_o)[y-F(x_o)]; \tag{2.3}$$

g is an element of X.

Putting $d = x-x_o$, we define

$$\tilde{Q}(d) = \langle y-F(x_o)-P(x_o)d, y-F(x_o)-P(x_o)d \rangle;$$

$\tilde{Q}$ is bilinear in d and will be a good approximation to Q in a sufficiently small neighborhood of x_o, i.e., for small d. We shall call a set of points

$\left\{ x : \tilde{Q}(d) \leq \mu \right\}$ for fixed μ an ellipsoid.

For $\lambda \geq 0$, we write

$$(A + \lambda I)d = g. \tag{2.4}$$

We note that when $\lambda = 0$, (2.4) is just the usual "normal equation" of least squares, and $\hat{d} = A^{-1}g$ determines the unique minimum of $\tilde{Q}(d)$. Theorem 1 suggests the way in which (2.4) can be used to construct a method for minimizing $Q(x)$. The proof is similar to the stronger version of Marguardt's Theorem 1 contained in Meeter [7]. As this paper was being written, we discovered essentially the same result, with a different proof and motivation, in Balakrishnan [1], pp. 160-161.

THEOREM 1. <u>Let</u> d_o <u>be the solution of (2.4) for a fixed value of</u> $\lambda \geq 0$. <u>Then</u> d_o <u>determines the minimum of</u> $\tilde{Q}$ <u>everywhere except over the interior of the ellipsoid</u> Ω: $\left\{x:\tilde{Q}(d) \leq \tilde{Q}(d_o)\right\}$. <u>In particular</u>, d_o <u>minimizes</u> Q <u>uniquely over the sphere</u> Φ <u>centered at</u> x_o <u>with radius</u> $\| d_o \|$.

PROOF. We first observe that, in the usual least squares manner, Ω can be expressed as the set of all d such that

$$< A(d-\hat{d}), d-\hat{d} > \leq \omega, \tag{2.5}$$

where $\omega = \tilde{Q}(d_o) - \| y-F(x_o)-P(x_o)\hat{d} \|^2$.

The first assertion of the theorem is obvious, since equality holds in (2.5) only for boundary points of Ω, and from the original definition of Ω we see that d_o is a boundary point.

More importantly for the purposes of the iterative method, we show that the sphere Φ centered at x_o is included in the region in which $\tilde{Q}$ is minimized by d_o. There are many ways to do this. One could adapt the proof found in Morrison [9], or use the argument found in Meeter [7]. Here we show that the intersection of Φ and Ω consists of the point d_o.

First we note that from (2.4) and the definition of $\hat{d}$,

$$A(d_o-\hat{d}) = (g-\lambda d_o) - AA^{-1}g = -\lambda d_o \tag{2.6}$$

Also, for all d in Φ,

$$< d_o,d>/\|d_o\|^2 \leq <d_o,d>/\|d\|\ \|d_o\| \leq 1 = <d_o,d_o>/\|d_o\|^2,$$

so that we can say

$$< d_o,d> \leq < d_o,d_o>, \tag{2.7}$$

with equality holding only if $d = d_o$. Now take some point $\bar{d}$ in Φ. We have

$$< A(\bar{d}-\hat{d}),\ \bar{d}-\hat{d}> = < A(\bar{d}-d_o+d_o-\hat{d}),\ \bar{d}-d_o+d_o-\hat{d}>$$

$$= < A(\bar{d}-d_o),\ \bar{d}-d_o> + 2< A(d_o-\hat{d}),\ \bar{d}-d_o> + \omega .$$

The first term is non-negative. From (2.6), the second term is

$$2<-\lambda d_o,\ \bar{d}-d_o> = 2\lambda[< d_o,\ d_o> - < d_o,\ \bar{d}>].$$

Applying (2.7), the second term is also non-negative, and zero only if $d_o = \bar{d}$. (If $\lambda = 0$, $d_o = d$, Ω is a point and the theorem is a trivial result). Thus, for $d \in \Phi$,

$$< A(d-\hat{d}),\ d-\hat{d} > > \omega$$

unless $d = d_o$, which proves that Φ and Ω have only the point d_o in common. This implies that the unique minimum of $\acute{\tilde{Q}}$ over Φ is attained at d_o.

Now let us regard d_o, the solution of (2.4), as a function of λ. A way of controlling the length of d_o is given by

THEOREM 2. <u>$\|d_o(\lambda)\|$ is a continuous, strictly decreasing function of λ such that $\|d_o(\lambda)\| \to 0$ as $\lambda \to \infty$.</u>

PROOF. Let $d_1 = d_o(\lambda_1)$, $d_2 = d_o(\lambda_2)$, $\lambda_1-\lambda_2 = \gamma$.

Then, from (2.4),

$$(A+\lambda_1 I)d_1 = (A+\lambda_2 I)d_2 , \tag{2.8}$$

or $$A(d_1-d_2) + \lambda_2(d_1-d_2) = -\gamma d_1.$$

Thus, taking norms,

$$\| d_1\|^2 = \gamma^{-2}\| (A+\lambda_2 I)(d_1-d_2)\|^2 . \tag{2.9}$$

Similarly, we can obtain

$$\| d_2\|^2 = \gamma^{-2}\| (A+\lambda_1 I)(d_1-d_2)\|^2$$

Then

$$\| d_2\|^2-\| d_1\|^2 = \gamma^{-2}\Big[< (A+\lambda_1 I)(d_1-d_2), (A+\lambda_1 I)(d_1-d_2)$$
$$- (A+\lambda_2 I)(d_1-d_2), (A+\lambda_2 I)(d_1-d_2)> \Big]$$
$$= \gamma^{-2}\Big[(\lambda_1{}^2-\lambda_2{}^2) < d_1-d_2,d_1-d_2 > + 2(\lambda_1-\lambda_2) < A(d_1-d_2), d_1-d_2> \Big].$$

The expression in square brackets has the same sign as $\lambda_1-\lambda_2$, and is zero only if $d_1 = d_2$, so that we can assert $\| d_1\|^2 < \| d_2\|^2$ when $\lambda_1 > \lambda_2$, i.e., $\| d_o(\lambda)\|$ is strictly decreasing.

Using (2.8), and holding λ_2 fixed, with $\lambda_1 > \lambda_2$,

$$\| d_1\|^2 \leq \gamma^{-2}\| (A+\lambda_2 I)\|^2 \| d_1-d_2\|^2$$
$$\leq \gamma^{-2}\| (A+\lambda_2 I)\|^2 \Big[\| d_1\|^2+\| d_2\|^2 \Big]$$
$$< 2\gamma^{-2}\| (A+\lambda_2 I)\|^2 \| d_2 \|^2.$$

Letting $\lambda_1 \to \infty$ shows that $\| d_o(\lambda)\| \to 0$.

To show that $\| d_o(\lambda)\|$ is a continuous function of λ, from (2.8) we obtain

$$d_1-d_2 = -(\lambda_1-\lambda_2)(A+\lambda_1 I)^{-1}d_2,$$

thus

$$\| d_1-d_2\| \leq |\lambda_1-\lambda_2| \; \| (A+\lambda_1 I)^{-1}\| \; \| d_2\| \; . \tag{2.10}$$

From (2.2) we know that $<(A+\lambda_1 I)x,x> \; \geq \; (\beta_2+\lambda_1) <x,x>$, for all $x \in X$, where β_2 is fixed and positive. Thus, as long as $\lambda_1 \geq -\epsilon > -\beta_2$, where ϵ is some constant > 0, $(A+\lambda_1 I)^{-1}$ exists and $\| (A+\lambda_1 I)^{-1}\| \leq \frac{1}{\epsilon}$. Holding λ_2 fixed, $\lambda_2 \geq 0$, we see from (2.10) that $\| d_1-d_2\| \to 0$ as $\lambda_1 \to \lambda_2$, hence $\| d_1\| \to \| d_2\|$ as $\lambda_1 \to \lambda_2$, proving $\| d_o(\lambda)\|$ is continuous at λ_2.

The connection between this method and the gradient method, and a means of controlling the orientation of the correction vector, is established by

THEOREM 3. <u>The angle α between d_o and g is a continuous, strictly decreasing function of λ. As $\lambda \to \infty$, $\alpha \to 0$, and d_o rotates toward g.</u>

PROOF. We will make frequent use of the fact that the operator $A+\lambda I$ and its inverse are linear, self-adjoint, strictly positive, and bounded. For convenience, we will show $\cos^2\alpha$ is increasing, where

$$\cos^2\alpha = \; <g,(A+\lambda I)^{-1}g>^2/\| g \|^2 \| (A+\lambda I)^{-1}g\|^2 .$$

The denominator of $\cos \alpha$ is a continuous function of λ, from Theorem 2. As for the numerator, using the notation of Theorem 2,

$$|<g,d_1> - <g,d_2>| = |<g,d_1-d_2>| \leq \| g\| \; \| d_1-d_2\| ,$$

and holding λ_2 fixed, we know $\| d_1-d_2\| \to 0$ as $\lambda_1 \to \lambda_2$.

Hence $\cos \alpha$ is the ratio of two continuous functions so α is a continuous function of λ, $\lambda \geq 0$.

Now we examine the behavior of $\cos^2\alpha$ in a small neighborhood of some fixed value of λ. For $|\gamma|$ sufficiently small, $A+(\lambda+\gamma)I$ is a strictly positive self-adjoint operator. Writing $B = A+\lambda I$,

$$(B+\gamma I)^{-1} = (B^{-1}(B+\gamma I))^{-1} B^{-1} = (I+\gamma B^{-1})^{-1}B^{-1} .$$

We make the one-to-one transformation $g = Bz$, so that for small $|\gamma|$, fixed λ,

$$\cos^2\alpha = < Bz,(I+\gamma B^{-1})^{-1}z>^2/ < Bz,Bz> <(I+\gamma B^{-1})^{-1}z,\ (I+\gamma B^{-1})^{-1}z >. \tag{2.11}$$

We can choose $|\gamma|$ small enough to have $\|\gamma B^{-1}\| < 1$. Accordingly, the power series expansion of the operator $(I+\gamma B^{-1})^{-1}$ will be convergent. The numerator of (2.11) can be rewritten as

$$< Bz,(I-\gamma B^{-1}+\gamma^2B^{-2}-....)z>^2 = \left[< Bz,z> - \gamma < Bz,B^{-1}z> + \gamma^2< Bz,B^{-2}z> -....\right]^2$$

$$= < Bz,z>^2-2\gamma < Bz,z> < z,z> + o(\gamma) . \tag{2.12}$$

The remaining factors in (2.11) are written as

$$< Bz,Bz>^{-1} \left[< (I-\gamma B^{-1}+ \gamma^2B^{-2}-....)z,\ (I-\gamma B^{-1}+\gamma^2B^{-2}-....)z >\right]^{-1}$$

$$= < Bz,Bz>^{-1} \left[<z,z>-2\gamma < B^{-1}z,z> + o(\gamma) \right]^{-1}$$

$$= < Bz,Bz>^{-1}<z,z> \left[1-2\gamma < B^{-1}z,z> < z,z>^{-1}+ o(\gamma) \right]^{-1}$$

$$= < Bz,Bz>^{-1}<z,z> \left[1+2\gamma < B^{-1}z,z><z,z>^{-1}+ o(\gamma)\right] \tag{2.13}$$

Multiplying (2.12) and (2.13), we obtain

$$\cos^2\alpha = \|Bz\|^{-2}\Big(<Bz,z>^2<z,z>+2\gamma[<B^{-1}z,z><Bz,z>^2-<z,z>^2<Bz,z>]\Big)+ o(\gamma).$$

Thus to show $\cos^2\alpha$ is strictly increasing, we need only to show that the term in square brackets is strictly positive. That is, we must show that

$$< B^{-1}z,z> < Bz,z> \ > \ < z,z>^2 . \tag{2.14}$$

From Schwarz's inequality, for any w,

$$< w,s> < Bw,Bw> \ > \ < w,Bw>^2, \tag{2.15}$$

unless w and Bw are linearly dependent. But if w and Bw are linearly dependent then $B = \mu I$, implying either $\cos^2\alpha \equiv 1$, or $w = 0$. But we will require $w = C^{-1}z$, where $C^2 = B$. Since $z = B^{-1}g$, $w \neq 0$ because $g \neq 0$. (If $g \neq 0$ we have achieved a relative minimum of Q, and iteration ceases.) The substitution $w = C^{-1}z$ reduces (2.15) to (2.14). Thus, for all $\lambda \geq 0$, $\cos\alpha$ is a continuous, strictly increasing function of λ.

A similar technique shows that $\cos\alpha \to 1$ as $\lambda \to \infty$. Briefly, for $\lambda > 0$, we rewrite $\cos^2\alpha$ as

$$\cos^2\alpha = \langle g,(I+\lambda^{-1}A)^{-1}g\rangle^2 / \langle g,g\rangle \langle (I+\lambda^{-1}A)^{-1}g,\ (I+\lambda^{-1}A)^{-1}g\rangle .$$

For λ sufficiently large, we can again use the power series expansion to obtain

$$\cos^2\alpha = \frac{\langle g,g\rangle^2 - 2\lambda^{-1}\langle Ag,g\rangle \langle g,g\rangle + o(\lambda^{-1})}{\langle g,g\rangle\ [\langle g,g\rangle - 2\lambda^{-1}\langle Ag,g\rangle + o(\lambda^{-1})]} ,$$

which shows that $\cos\alpha \to 1$ as $\lambda \to \infty$. Since g is fixed, d_o rotates towards g as $\lambda \to \infty$.

A convergence proof for the method may be obtained as was done by Tornheim [10] for Euclidean n-space. Let $S = \left\{x\varepsilon X : Q(x) \leq x_o\right\}$.

THEOREM 4. <u>Suppose $Q(x)$ has a second (Gateaux) derivative $Q''(x,h,h)$, and suppose there exists $\rho_o > o$ such that $|Q''(x,h,h)| \leq \|h\|^2/\rho_o$ for all $x\varepsilon S$, $h\varepsilon X$. Then it is possible to choose a sequence λ_n such that $Q(x_{n+1})$ converges downward to a limit, where $x_{n+1}-x_n = d_n + (A_n + \lambda_n I)^{-1}g_n$, $n = 0,1,\ldots$.</u>

PROOF. The above conditions are sufficient to insure that by correcting x_n with a vector μg_n, $\mu>0$, we can by proper choice of μ have $Q(x_n + \mu g) < Q(x_n)$, $\|g_n\| \neq 0$. See Goldstein [2]. Since d_n has a positive projection on g_n, and $Q(x)$ is continuous, Theorems 2 and 3 indicate that it will always be possible to choose λ_n sufficiently large that $Q(x_{n+1})<Q(x_n)$. Since $Q(x)$ is bounded below, the sequence $Q\left(x(\lambda_n)\right)$ converges downward to some limit.

3. <u>The Connection with Newton's Method</u>. Although the method as developed in Section 2 is actually a type of interpolation between the Taylor Series or Gauss Method and the gradient method, we can also regard it as connecting Newton's Method and the gradient or steepest descent method.

Suppose now that F is a nonlinear operator $F:X\to X$ where X is a real Hilbert space. If, for some $x\varepsilon X$, a solution to the equation

$$F(x) = 0 \tag{3.1}$$

exists, and F has a Frechet derivative $P(x,h)$, Newton's method for solving (3.1) is written as the sequence

$$x_{n+1} = x_n - P(x_n)^{-1}F(x_n), \tag{3.2}$$

$n = 0,1\ldots$, derived by equating to the zero vector a linear approximation to F at x_n. See Kantorovich [3] or Moore [8]. If we put $d_n = x_{n+1} - x_n$, and define again the linear self-adjoint operator $A = P^*(x_n)P(x_n)$, we obtain from (3.2)

$$-Ad_n = P^*(x_n)F(x_n). \tag{3.3}$$

On the other hand, in order to solve (3.1) by the gradient method, we might seek to solve the functional equation

$$f(x) = \langle F(x),F(x)\rangle = 0$$

by making our correction g_n to x_n proportional to the negative gradient of $f(x)$, or

$$g_n \propto -2P^*(x_n)F(x_n),$$

which would mean that our method would determine d_n from

$$(A+\lambda I)d_n = g_n,$$

as before. Since both the Taylor Series-Gauss and Newton methods begin with the same linear approximation to a nonlinear operator F, perhaps the designation "Newton-Gradient" can be justified on the grounds of euphony.

At a later date we hope to be able to investigate the more difficult questions of regions and speeds of convergence, and present examples.

4. Acknowledgment. The preliminary work for this paper was performed while the author was the recipient of a Research Council Grant of the Florida State University.

5. References

[1] A. V. Balakrishnan, "Optimal control problems in Banach spaces," J. SIAM Control, 3, (1965), pp. 152-180.

[2] A. A. Goldstein, "Minimizing functionals on Hilbert space," in Computing Methods in Optimization Problems, A. V. Balakrishnan and L. W. Neustadt, eds., Academic Press, New York, 1964, pp. 159-165.

[3] L. V. Kantorovich and G. P. Akilov, Functional Analysis in Normed Spaces, Pergamon Press, New York, 1964.

[4] K. Levenberg, "A method for the solution of certain nonlinear problems in least squares," Quart. Appl. Math., 2, (1944), pp. 164-168.

[5] D. L. Marquardt, "An algorithm for least squares estimation of nonlinear parameters," J. Soc. Indust. Appl. Math., 2, (1963), pp. 431-441.

[6] D. A. Meeter, "Problems in the analysis of nonlinear models by least squares," Ph.D. thesis, University of Wisconsin, Madison, 1964.

[7] D. A. Meeter, "On a theorem in nonlinear least squares," J. Soc. Indust. Appl. Math., (to appear).

[8] R. H. Moore, "Newton's method and variations," in Nonlinear Integral Equations, P. M. Anselone, ed., University of Wisconsin Press, Madison, 1964.

[9] D. D. Morrison, "Methods for nonlinear least squares problems and convergence proofs," Proc. Jet Propulsion Lab. Seminar: Tracking Problems and Orbit Determination, (1960), pp. 1-9.

[10] L. Tornheim, "Convergence in nonlinear regression," Technometrics, 5, (1963), pp. 513-514.