POLYNOMIAL LEAST SQUARES APPROXIMATIONS WITH ILL-CONDITIONED BASES

by

G.R. Richter and W.L. Steiger

DCS-TR-#47

August 1976

DCS Technical Report #47

Department of Computer Science
Rutgers University
New Brunswick, N.J.

ABSTRACT

In this note we point out that polynomial least squares approximations may be unstable in coefficient space and stable in L_2. That is, an ill-conditioned basis may produce large errors in particular coefficients yet the L_2 error is small. An explanation of this phenomenon is offered.

Given $f \in L_2[a,b]$, the Hilbert space of real, square integrable functions on $[a,b]$ with inner product $\langle g,h\rangle = \int_a^b g(x)\, h(x)\, dx$, we seek the polynomial $P_n(x)$ which minimizes

$$\left\| f - P_n \right\| = \langle f - P_n,\ f - P_n\rangle^{1/2}$$

over the set of polynomials of degree $\le n$. Representing $P_n(x)$ as a linear combination

$$P_n(x) = \sum_{j=0}^{n} c_j \phi_j(x)$$

of a given set of basis functions $\phi_o, \ldots, \phi_n$ for the n^{th} degree polynomials, the optimal set of coefficients $c_o, \ldots, c_n$ is obtained by solving the normal equations

$$(*) \qquad E\, \underline{c} = \underline{d}$$

where $e_{ij} = \langle\phi_i,\phi_j\rangle$, $0 \le i,j \le n$, and $d_i = \langle f,\phi_i\rangle$, $0 \le i \le n$.

In practice, the normal equations are usually solved on a computer; this produces a coefficient vector $\hat{c}$ and a corresponding approximation $\hat{P}_n(x)$ which are contaminated as a result of round-off error. If the basis $\{\phi_o, \ldots, \phi_n\}$ is nearly linearly dependent, which occurs, for example, for the monomials $\phi_j(x) = x^j$, E will be an ill-conditioned matrix and the computation of $\hat{c}$ quite unstable. In particular, if the condition number of E substantially exceeds the number of digits of computational precision, one may anticipate a total breakdown of significance in $\hat{c}$. For this reason, there has been a tendency to avoid the expedient monomial basis and, where feasible, use an orthogonal basis instead.

This notwithstanding, we wish to point out an observation which may not be of wide currency and indeed, came as a surprise to us.

IT IS OFTEN THE CASE, IN THE PRESENCE OF SEVERELY ILL-CONDITIONED NORMAL EQUATIONS, THAT $\left\| f - \hat{P}_n \right\|$ *IS SMALL DESPITE GROSS IMPRECISION IN* $\hat{c}$.

Example: $f = \dfrac{1}{1 + 9x^2} \in L_2\,[-1,1]$ and $n = 18$. The even coefficients of $\underline{c}$ and $\hat{\underline{c}}$, the latter computed in single precision on a PDP 10 computer, are as follows:

i	c_i	$\hat{c}_i$
0	.998	.988
2	-8.572	-7.312
4	60.916	34.385
6	-309.765	-88.172
8	1051.240	104.240
10	-2323.137	-14.305
12	3279.462	-64.953
14	-2840.825	6.125
16	1372.625	58.050
18	-282.845	-28.953

The true c_i were computed by solving (*) in the rationals (with no round-off) while the $\hat{c}_i$ were computed using Gaussian elimination with partial pivoting, then backsolving. In both cases, the odd coefficients vanish since f is even. The sup-norm condition number of E is 2.0958×10^{13}, again computed in the rationals, so E is very ill-conditioned and indeed, $\hat{\underline{c}}$ is very far from $\underline{c}$. On the other hand, $\left\|f - \hat{P}_{18}\right\| = .0085$ while $\left\|f - P_{18}\right\| = .0012$. Similar results occurred when f was taken to be a discontinuous function or various polynomials, and n ranged as high as 64.

Least squares approximations may us be quite robus even in the face of ill-conditioned nmal equations that a... ng from an almost dependent basis. That this is possible is not in itself a startling revelation. Due to the near dependence of the basis, $P_n - \hat{P}_n$ might well have large coefficients and a small L_2 norm. Indeed, the set of coefficients in R^{n+1} that gives rise to approximations within $\epsilon > 0$ (in L_2) of P_n is an ellipsoid, $S_n(\epsilon) = \{\underline{t} \in R^{n+1} : (\underline{t} - \underline{c})^T E (\underline{t} - \underline{c}) < \epsilon\}$, where E is the symmetric, positive definite matrix of the normal equations and $\underline{c}$, the coefficients of P_n. In the monomial basis, successive axes of $S_n(\epsilon)$ roughly double in length, thus allowing $\left|c_j - \hat{c}_j\right|$ to increase as j does, for "within ϵ" approximations. In this way least squares approximations may have large errors in particular coefficients while having small L_2 error. What is surprising, however, is that roundoff errors in the solution of ill-conditioned normal equations actually tend to favor this situation.

To explain this, consider more closely Gaussian elimination as applied to the normal equations of the example; here $e_{ij} = 2/(i + j - 1)$ if $i + j$ is even and 0 otherwise. At step k in the elimination write $\hat{e}_{kk}^{(k)}$ for the pivot element used in the k^{th} column of E, while $e_{kk}^{(k)}$ is the exact value of this quantity. In general, the larger the index k, the closer $e_{kk}^{(k)}$ is to zero -- a manifestation of the ill-conditioning in E. During elimination, the pivots arise essentially as algebraic sums of numbers, the largest of which is roughly of order one. Therefore, the absolute accuracy in the computed pivots cannot be expected to exceed the unit of roundoff error δ on the computer used for the computation. The result is that the first few computed pivots are quite accurate, while those for which $e_{kk}^{(k)} < \delta$ contain no precision whatever; in fact, the latter are of magnitude $\approx \delta$ and do not decay at all with increasing k. For the preceding example, the situation is as depicted below.

k	$e_{kk}^{(k)}$	$\hat{e}_{kk}^{(k)}$	Exponent
7	5.3043	5.3043	-4
8	1.6809	1.6809	-4
9	-2.9399	-2.9399	-5
10	-4.9213	-4.9189	-6
11	-9.9878	-10.0047	-7
12	-2.7050	-2.7811	-7
13	-4.3464	-3.4448	-8
14	-1.1245	-1.6442	-8
15	-1.1778	-6.0617	-9
16	-2.5789	-79.2683	-10
17	-2.8809	-80.0467	-11
18	-5.6747	3762.2723	-12

Again, true values were computed in the rationals, the first few k omitted for brevity.

Looking now at the first step of the backsubstitution, we see that the result of the magnitude distortion in $\hat{e}_{nn}^{(n)}$ is to make the magnitude of $\hat{c}_n$ smaller than that of c_n. Likewise, all $\hat{c}_k$ for which the corresponding pivots $\hat{e}_{kk}^{(k)}$ are distorted will tend to be of smaller magnitude than their true values. Moreover, the effect of these sub-magnitude $\hat{c}_k$'s is to dampen the components of the corresponding basis functions $\phi_k(x)$ in the computed approximation $\hat{P}_n(x)$, which is borne out by the computational results in the preceding example. In essence, $\hat{P}_n(x)$ is a smoothed approximation to f for large n, which is close in the L_2 sense to that $P_k(x)$ for which $\left|e_{kk}^{(k)}\right| \approx \delta$.

Thus the monomial basis does not lead to instability (i.e. unboundedness) in $\|f - P_n\|$ as $n \to \infty$ and it is this quantity which is often of greatest importance from a practical standpoint. We do not suggest that the monomial basis should be used for polynomial least squares approximation problems, although in cases where f can be approximated well by a low degree polynomial a small error $\|f - \hat{P}_n\|$ is likely to occur. Rather, we wish to point out what amounts to an interesting dichotomy between stability in the coefficient sense and stability in the L_2 sense.