NUMERICAL METHODS FOR INVERSE EIGENVALUE PROBLEMS IN ALGEBRAIC CONTROL THEORY

J. Kautsky, N.K. Nichols and P. Van Dooren

In this talk we outline three numerical methods for solving the following problem (details are to be reported elsewhere, see also [2]):

Given n linear subspaces $S_j \subset E_n$ in the n-dimensional real vector space choose one vector $\underline{x}_j \in S_j$, $j = 1,2,\ldots,n$ in each so that these n vectors $\underline{x}_1,\underline{x}_2,\ldots,\underline{x}_n$ are as orthogonal as possible.

Problems of this kind arise, for example, in algebraic control theory when, given an $n\times n$ matrix A , an $n\times m$ matrix B of rank m and numbers $\lambda_1,\lambda_2,\ldots,\lambda_n$ we seek an $m\times n$ matrix F such that the eigenvalues of the matrix $A+BF$ are the given numbers $\lambda_1,\ldots,\lambda_n$. For $m>1$ there may be many solutions F and it is then desirable to construct that F for which the eigenvalues λ_j of $A+BF$ are least sensitive to perturbations. This sensitivity is proportional to the condition numbers c_j (see [4]) of eigenvalues λ_j given by $c_j = \|\underline{x}_j\| \, \|\underline{y}_j\|$ where the right eigenvectors $\underline{x}_j$ satisfy

$$(1) \qquad (A - \lambda_j I + BF)\underline{x}_j = \underline{0}$$

and $\underline{y}_j$ are the left eigenvectors given by

$$\underline{y}_j = X^{-T}\underline{e}_j \quad , \quad j = 1,2,\ldots,n \ ,$$

where $X = (\underline{x}_1,\underline{x}_2,\ldots,\underline{x}_n)$. As $c_j \geq 1$ with equality (for all $j = 1,2,\ldots,n$) occuring iff the columns $\underline{x}_j$ of X are orthogonal there are many measures of the vector $\underline{c} = (c_1,\ldots,c_n)^T$ which can be minimized to express mathematically the intuitive notion, used above, of vectors $\underline{x}_j$ being "as orthogonal as possible".

The subspaces S_j for the eigenvectors in (1) are easily constructed for the given data as right nullspaces of $U_1^T(A-\lambda_j I)$ where the columns of U_1 are any (preferably orthogonal) basis for the left nullspace of B.

We comment that for the construction of F as well as for other properties important from the control theory view-point it is important that the matrix X is well conditioned ([1]). Therefore, and also to motivate our numerical methods, we now list some relations between norms of the vector of sensitivities $\underline{c} = (c_1, c_2, \ldots, c_n)^T$ and condition numbers of X. Without loss of generality (the c_j's are independent of scaling of $\underline{x}_j$'s) we assume that the columns of X are normalized, so that $\|X\underline{e}_j\| = \|\underline{x}_j\| = 1$. Let $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$ be a diagonal matrix. We have ($\|A\|_F = (\sum_{i,j} (\underline{e}_i^T A \underline{e}_j)^2)^{\frac{1}{2}}$)

(2) $$\mathrm{cond}_F(X^{-T}D) = \mathrm{cond}_F(XD^{-1}) = (\sum_i d_i^{-2})^{\frac{1}{2}} (\sum_i d_i^2 c_i^2)^{\frac{1}{2}},$$

(3) $$\|\underline{c}\|_2 = \|X^{-1}\|_F = n^{-\frac{1}{2}} \mathrm{cond}_F(X),$$

(4) $$\|\underline{c}\|_1 = \mathrm{cond}_F(XD^{-1}) \text{ if } d_i = c_i^{-\frac{1}{2}},$$

(5) $$\|\underline{c}\|_\infty \le \mathrm{cond}_2(X) \le \mathrm{cond}_F(X).$$

Also, if $\hat{X} = (\hat{\underline{x}}_1, \hat{\underline{x}}_2, \ldots, \hat{\underline{x}}_n)$ is an orthogonal matrix $(\hat{X}^T\hat{X} = I_n)$ and the sines $\varphi_j = (1 - \hat{\underline{x}}_j^T \underline{x}_j)^{\frac{1}{2}}$ of the angles between vectors $\hat{\underline{x}}_j$ and $\underline{x}_j$ are sufficiently small then

(6) $$\|\underline{c}\|_2 \le \sqrt{n} + \rho(1-\rho)$$

where $\rho = (2 \sum_j \varphi_j^2)^{\frac{1}{2}}$.

Note that (2) allows us to represent $\|D\underline{c}\|_2$ exactly by the Frobenius norm of the scaled inverse of X. Actually, weighted p-norms of $\underline{c}$ (including the uniform norm as a limit) may be expressed similarly if the scaling D is allowed to depend on $\underline{c}$ itself which can be achieved by our methods by adaptive iterations. One particular example is relation (4) where the scaling is the optimal scaling for minimizing both the Frobenius and spectral conditions of X. We see that various norms of $\underline{c}$, $\text{cond}_2(X)$ and $\text{cond}_F(X)$ are closely inter-related and that they all reach minimum iff the matrix X is orthogonal.

All our methods to find a suitable selection of vectors $\underline{x}_j \in \mathcal{S}$, $j = 1,2,\ldots, n$ are iterative. In the first *method 0* an elementary iteration consists of replacing one of the columns of X, say $\underline{x}_1$, by a vector from $\mathcal{S}_1$ which minimizes c_1. This new $\underline{x}_1$ is found easily as an orthogonal projection, into $\mathcal{S}_1$, of the vector $\underline{y}_1$ orthogonal to all other columns of X. Sweeps of n such iterations, replacing in turn each column of X are repeated until some general condition, say $\text{cond}_2(X)$, has hopefully settled to an acceptable minimum. However, convergence of this kind cannot be assured as improving one of the condition numbers may worsen the others and, indeed, numerical experiments confirmed this non-convergence, although good results were obtained by *method 0* asymptotically.

The second *method* 1 performs the same sweeps of elementary operations in which, however, the new column of X, say $\underline{x}_1$ is chosen to minimize $\|D\underline{c}\|_2$ (for some prescribed scaling D). As this is a global measure, independent of the updated vector $\underline{x}_1$, the convergence of the method is assured. Denoting by S_1 some orthogonal basis of $\mathcal{S}_1$ and $X_1 = (\underline{x}_2, \underline{x}_3, \ldots, \underline{x}_n)$ the elementary iteration comprises finding an m-vector $\underline{u}$ of unit length $\|\underline{u}\|_2 = 1$ such that $\|D(S_1\underline{u}, X_1)^{-1}\|_F$ is minimized. This is a non-linear constrained least square type problem which, however, can be solved explicitly essentially by three orthogonal decompositions.

Our third *method 2* is based on minimizing ρ in (6). Instead of X we aim to position an orthogonal matrix $\hat{X}$ in such a way that each of its columns $\hat{\underline{x}}_j$ is close to the corresponding subspace S_j. The result is then obtained by projecting $\hat{\underline{x}}_j$ into S_j. The positioning of $\hat{X}$ is done by sweeps of $n(n-1)/2$ rotations in planes determined by pairs of vectors $\hat{\underline{x}}_j$, $\hat{\underline{x}}_k$, $j \neq k$; the angles of which are chosen to minimize $\varphi_j^2 + \varphi_k^2$ where φ_j is now the sine of the angle between $\hat{\underline{x}}_j$ and S_j. These elementary rotations are similar to those occurring in the Jacobi method for calculating eigenvalues and can be obtained explicitly. The calculations involve scalar products of vectors of length m which, for $m > n-m$, may be replaced by vectors of length $n-m$ by using orthogonal complements of subspaces S_j (*method 3*). As the determination of the rotation requires less effort than the actual update of the matrix $\hat{X}$ a threshold technique can be employed to increase efficiency and to ensure convergence.

We note that a similar method was proposed by Klein and Moore [3] where, essentially, the objective of our *method* 1 was combined with the plane rotations technique of *method 2/3*. In this case, however, the optimal rotation could not be obtained explicitly so that another iterative process had to be performed to complete each elementary iteration.

It is interesting to note that an *a priori* lower bound for the obtainable conditioning can be derived; indeed, for any X such that $X\underline{e}_j \in S_j$, $\|X\underline{e}_j\| = 1$ we have

$$\mathrm{cond}_2(X) \geq n^{-\frac{1}{2}}\,\mathrm{cond}_2(S)$$

where $S = (S_1, S_2, \ldots, S_n)$ is a matrix of combined orthonormal bases S_j of subspaces S_j. Although this bound is not sharp it provides a useful information on *a priori* poorly conditioned situations as well as indication how to proceed in the "inverse inverse" eigenvalue problem of the algebraic control theory: how to choose $\lambda_1,\ldots,\lambda_n$ to achieve a robust result.

Finally we wish to comment that similar approaches and numerical methods can be applied to other problems in control theory, for example output feedback problems and state feedback problems for descriptor systems.

REFERENCES

[1] J. Kautsky, N.K. Nichols, P. van Dooren and L. Fletcher, *Numerical methods for robust eigenstructure assignment in control system and design*, Proc. of Workship on "Numerical Treatment of Inverse Problems for Differential and Integral Equations", Heidelberg, 1982, Birkhauser, Boston.

[2] J. Kautsky and N.K. Nichols, *Robust Eigenstructure Assignment in State Feedback Control*, Numerical Analysis Rpt. NA/2/83, University of Reading.

[3] G. Klein and B.C. Moore, *Eigenvalue-generalized eigenvector assignment with state feedback*, IEE Trans. Auto. Control AC-22, 140-141 (1977).

[4] J.H. Wilkinson, *The algebraic eigenvalue problem*, Oxford University Press, Oxford (1965).

J. Kautsky
School of Mathematical Sciences
Flinders University
Bedford Park, S.A. 5042
AUSTRALIA.

K.N. Nichols
Department of Mathematics
University of Reading
Whiteknights, Berks., RG62AX
U.K.

P. Van Dooren
Philips Research Laboratory,
Av. van Beceleare, 2, Box 8
B-1170 Brussels
BELGIUM