Math221 Fall 2009 Homework # 8 Solutions

Instructor: James W. Demmel

November 5, 2009

Prob 4.6

Consider the Schur decompositions $A = Q_A T_A Q_A^*$ and $B = Q_B T_B Q_B^*$. Then $AX - XB = C$ implies $Q_A T_A Q_A^* X - X Q_B T_B Q_B^* = C$. Premultiplying both sides by Q_A^*, postmultiplying both sides by Q_B, and substituting $X = Q_A Y Q_B^*$ and $C = Q_A C' Q_B^*$ yields $T_A Y - Y T_B = C'$.

This is triangular system of equations for Y in (a slight) disguise: We solve for the entries of Y one at a time, starting with Y_{m1}, continuing up the first column $Y_{m-1,1}, ..., Y_{11}$, and then the other columns of Y from left to right, again solving for each one from bottom to top. Equating i, j entries on both sides of $T_A Y - Y T_B = C'$ yields

$$\sum_{k=i}^{m} T_{A,ik} Y_{kj} - \sum_{k=1}^{j} Y_{ik} T_{B,kj} = C'_{ij}$$

Solving for Y_{ij} yields

$$Y_{ij} = \frac{C'_{ij} - \sum_{k=i+1}^{m} T_{A,ik} Y_{kj} + \sum_{k=1}^{j-1} Y_{ik} T_{B,kj}}{T_{A,ii} - T_{B,jj}}$$

Note that in the numerator, only terms involving entries of Y in the same column as Y_{ij} but in later rows ($T_{A,ik} Y_{kj}$) or in previous columns ($Y_{ik} T_{B,kj}$) are involved. So by computing the entries of Y in the order above, they can be solved for one at a time.

To be able to solve with arbitrary C, we also require that the denominator $T_{A,ii} - T_{B,jj}$ is never zero, for any i and j. This is equivalent to requiring that A and B have no common eigenvalues.

Once we have Y we compute $X = Q_A Y Q_B^*$.

In practice, if A and B are real, we can use real Schur form, and instead of a formula for each Y_{ij} as above, we can get a similar formula, or a 2-by-2 linear system, or a 4-by-4 linear system, depending on the sizes of the diagonal block encountered along the diagonals of T_A and T_B.

Prob 4.7

Let $T = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$, $M = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, $S = \begin{bmatrix} I & R \\ 0 & I \end{bmatrix}$. Then $S^{-1} = \begin{bmatrix} I & -R \\ 0 & I \end{bmatrix}$. $S^{-1} T S = M$ implies $AR - RB = -C$ which we solve as described in the previous question.

Prob 4.8

The desired similarity is as follows (where I_k is a k-by-k identity matrix):

$$\begin{bmatrix} I_m & -A \\ 0 & I_n \end{bmatrix} \cdot \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} \cdot \begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}$$

Each similar matrix is block triangular, and so its eigenvalues are the eigenvalues of the blocks, namely zeros along with the eigenvalues of AB and BA.

Prob 4.11

- Part 1: Changing x to αx and y to βy for nonzero scalars α and β does not change $(\alpha x)(\bar{\beta} y^*)/((\bar{\beta} y^*)(\alpha x)) = xy^*/y^*x$, since α and $\bar{\beta}$ cancel.
- Part 2: $P^2 = (xy^*xy^*)/(y^*x)^2 = (x(y^*x)y^*)/(y^*x)^2 = (xy^*)/(y^*x) = P$.
- Part 3: $AP = Axy^*/(y^*x) = \lambda xy^*/(y^*x) = \lambda P$, and $PA = xy^*A/(y^*x) = x\lambda y^*/(y^*x) = \lambda P$.
- Part 4: Assume without loss of generality that $\|x\|_2 = \|y^*\|_2 = 1$. Then $\|P\|_2 = \|xy^*\|_2/|y^*x| = \|x\|_2\|y^*\|_2/|y^*x| = 1/|y^*x|$. This (and Prob 4.13) rely on the fact that the 2-norm of a rank-1 matrix like xy^* is

$$\|xy^*\|_2 = \max_{\|z\|_2=1} |xy^*z| = \|x\|_2 \max_{\|z\|_2=1} |y^*z| = \|x\|_2\|y^*\|_2 \ .$$

Prob 4.13

If $r = Ax - \mu x$, $\|x\|_2 = 1$, then $\mu x = Ax - r = Ax - r(x^*x) = (A - rx^*)x \equiv (A + E)x$ where $\|E\|_2 = \|-rx^*\|_2 = \|r\|_2\|x^*\|_2 = \|r\|_2$.

Prob 4.16 Plug $x_i = g_i(t)$ for $i = 1, 2, 3$ into the polynomial matrix $F(x_1, x_2, x_3)$ to get $F(t) = F(g_1(t), g_2(t), g_3(t))$, a polynomial matrix in one variable t. Write $F(t) = \sum_{i=0}^{m} t^i A_i$ where A_i are constant matrices. The goal is to find real t where $\det F(t) = 0$, these values of t corresponding to intersections of the real parametric curve C given by $(g_1(t), g_2(t), g_3(t))$ and the real surface S given by $\det F(x_1, x_2, x_3) = 0$. Solving $\det F(t) = 0$ is a nonlinear eigenvalue problem, since higher powers of t than t^1 appear in $F(t)$. Following section 4.5.3 in the book, we build the following linear eigenvalue problem (a block companion matrix)

$$\hat{F}(t) = t \begin{bmatrix} A_m & & & \\ & I & & \\ & & \ddots & \\ & & & I \end{bmatrix} - \begin{bmatrix} -A_{m-1} & -A_{m-2} & \cdots & -A_0 \\ I & 0 & \cdots & 0 \\ & \ddots & \ddots & \\ & & I & 0 \end{bmatrix} \equiv Bt - C$$

and note that if $F(\tau)x = 0$ with $x \neq 0$, so that τ is an eigenvalue and x an eigenvector of $F(t)$, then $\hat{F}(\tau)\hat{x} = 0$, where $\hat{x} = [\tau^{m-1}x; \tau^{m-2}x; \cdots; \tau x; x]$, so τ is an eigenvalue and $\hat{x}$ an eigenvector of $\hat{F}(t)$. And it is easy to see that if z is any eigenvector of $\hat{F}(t)$, it must be of the form of $\hat{x}$, so the converse is true.

The eigenvalue problem for $Bt - C$ is equivalent to the problem for $tI - B^{-1}C$ if B is nonsingular, i.e. if A_m is nonsingular. So if A_m is nonsingular (and not too ill-conditioned!) we can solve the nonsymmetric eigenproblem for

$$B^{-1}C = \begin{bmatrix} -A_m^{-1}A_{m-1} & -A_m^{-1}A_{m-2} & \cdots & -A_m^{-1}A_0 \\ I & 0 & \cdots & 0 \\ & \ddots & \ddots & \\ & & I & 0 \end{bmatrix}$$

using QR iteration as discussed in class, computing the Schur form and then eigenvalues. Otherwise we need to treat the generalized nonsymmetric eigenproblem $tB - C$ using a generalization of QR iteration called QZ iteration (just eig(C,B) in Matlab) to compute the solution.

To see what happens when S and C do not intersect, consider the 1x1 problem where $F(x_1, x_2, x_3) = x_1$ (so S is the plane $x_1 = 0$) and $g_1(t) = 1 + t^2$, $g_2(t) = t$, $g_3(t) = 1$ (so C is a parabola). Then $F(t) = 1 + t^2$ which has no real solutions. Another example uses $F(x_1, x_2, x_3) = x_1$ again with $g_1(t) = 1$, $g_2(t) = t$ and $g_3(t) = 0$ (so C is a straight line parallel to S). Now $F(t) = 1$ has no solutions at all. (This would be described as an eigenvalue "at infinity" in the language of the Weierstrass Canonical Form, a generalization of Jordan form to $tB - C$ discussed in the book. Just as we do not compute the Jordan Form numerically, rather the Schur Form, we compute a generalized Schur Form instead of the Weierstrass form to discover such eigenvalues.)

To see what goes wrong when C lies entirely in S, consider $F(x_1, x_2, x_3) = x_1$ as above, and $g_1(t) = 0$, $g_2(t) = t$, $g_3(t) = 1 + t^2$, so C is a parabola lying entirely in the plane S. Then $F(t) = 0$ has *all* values of t as eigenvalues. (This means that $tB - C$ is a "singular pencil" in the language of the Kronecker Canonical Form, another generalization of Jordan Form mentioned in the book. Again, there is a stable numerical algorithm to analyze this situtation.)

For example 1 in the question in the book, $F(t) = \begin{bmatrix} 3t+3 & 0 \\ 0 & t+1 \end{bmatrix}$, so the eigenproblem is $t \cdot \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$, with characteristic polynomial $\det F(t) = 3(t+1)^2$.

For example 6 in the question in the book, $F(t) = \begin{bmatrix} 3t^2+3 & t^2 \\ 2+t^2 & t^2+1 \end{bmatrix}$ so the eigenproblem is $t \cdot \begin{bmatrix} 3 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & -3 & 0 \\ 0 & 0 & -2 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ with characteristic polynomial $\det F(t) = 2t^4 + 4t^2 + 3$.

For a complete solution of this problem, with Matlab implementation, see the class web page.