CS 383C
CAM 383C/M 383E Numerical Analysis: Linear Algebra Fall 2008

Solutions to Homework 3

Lecturer: Inderjit Dhillon Date Due: Sept 24, 2008

Keywords: *Singular Value Decomposition, Projectors, QR Factorization, Gram Schmidt Orthogonalization*

1. Problem 5.3

   (a) $U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$, $\Sigma = \begin{bmatrix} 10\sqrt{2} & 0 \\ 0 & 5\sqrt{2} \end{bmatrix}$, $V = \begin{bmatrix} -0.6 & 0.8 \\ 0.8 & 0.6 \end{bmatrix}$.

   (b)

   (c) $||A||_1 = \max_i \|a_i\|_1 = 16$, $||A||_2 = \sigma_1 = 10\sqrt{2}$, $||A||_\infty = \max_i \|a_i^*\|_1 = 15$.

   (d) $A^{-1} = V\Sigma^{-1}U^* = \begin{bmatrix} 0.05 & -0.11 \\ 0.10 & -0.20 \end{bmatrix}$

   (e) $\lambda_1 = 1.5 + \frac{\sqrt{391}}{2}i, \lambda = 1.5 - \frac{\sqrt{391}}{2}i$.

   (f) $det(A) = 100$

   (g) Area of an ellipse with semi-axis lengths a and b is given by πab. The matrix A maps the unit L_2 ball in $\mathbb{R}^2$ to an ellipse with semi-axis lengths σ_1, and σ_2. So, Area=100π.

2. Problem 5.4

   Given $A = U\Sigma V^T$. Let $B = \begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix}$, and let $\boldsymbol{z} = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix}$ be an eigenvector of matrix B. Thus,

   $$B\boldsymbol{z} = \lambda \boldsymbol{z} \quad \Longrightarrow \quad A^*\boldsymbol{y} = \lambda \boldsymbol{x} \quad \& \quad A\boldsymbol{x} = \lambda \boldsymbol{y}.$$

   Now $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$, where σ_i is the i-th singular value and $\boldsymbol{u}_i$ is the i-th column vector of matrix U. Similarly, $A^*\boldsymbol{u}_i = \sigma_i \boldsymbol{v}_i$. Therefore, $\boldsymbol{z} = \begin{bmatrix} \boldsymbol{v_i} \\ \boldsymbol{u_i} \end{bmatrix}$ is seen to be an eigenvector of B with eigenvalue σ_i. Normalize $\boldsymbol{z}$ to have $\|\boldsymbol{z}\|_2 = 1$ by dividing by $\sqrt{2}$ (note that $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are orthonormal). Similarly, $\frac{1}{\sqrt{2}} \begin{bmatrix} \boldsymbol{v_i} \\ -\boldsymbol{u_i} \end{bmatrix}$ is an eigenvector of matrix B with eigenvalue $-\sigma_i$. Thus, the eigenvalue decomposition of matrix B is given by:

   $$B = \frac{1}{2} \begin{bmatrix} V & V \\ U & -U \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{bmatrix} \begin{bmatrix} V^* & U^* \\ V^* & -U^* \end{bmatrix}$$

   You can check that the matrix $Q = \frac{1}{\sqrt{2}} \begin{bmatrix} V & V \\ U & -U \end{bmatrix}$ is orthonormal.

3. Problem 6.2

   If we denote the coordinates of vector x by x_j (where $j = 1, \ldots, m$), then the j^{th} coordinate of Ex is $\frac{1}{2}(x_j + x_{m+1-j})$. Therefore the j^th coordinate of E^2x is $\frac{1}{2}\left(\frac{1}{2}(x_j + x_{m+1-j}) + \frac{1}{2}(x_{m+1-j} + x_j)\right) = Ex$, meaning that E is a projector. The elements of (the matrix of) E are the coordinates of the Ee_i vectors, that is we have $\frac{1}{2}$ on both the diagonal and the anti-diagonal, and 0 elsewhere. The only exception is the element at intersection of the diagonal and anti-diagonal odd m values, where a 1 will show up. Since $E = E^*$, the projector is orthogonal (by Theorem 6.1 of the Textbook).

4. Problem 7.5

   (a) $\Longrightarrow$ Follows directly from part (b)

   $\Longleftarrow$ Follows directly from Theorem 7.2

   (b) **Claim**: rank(A)$\geq k$, where $0 \leq k \leq n$ (Note that although the Problem asks for $0 \leq k < n$ only, we prove the claim for a more general case which can be used for solving part (a))
   Proof: We prove by induction on the number of columns in matrix A. Let $A = [a_1, a_2, \ldots, a_n]$, $\hat{Q} = [q_1, q_2, \ldots, q_n]$, and r_{ij} is the $(i, j) - th$ element of $\hat{R}$.
   Base Case ($n = 1$): In this case $A = a_1 = q_1 r_{11}$. Thus clearly rank(A)$= 1$ if $r_{11} \neq 0$ else $A = 0$, thus rank(A)$= 0$.
   Induction step: Lets assume the claim for $n = p-1$. Now $A = [B\ a_p]$, where B is a $m \times (p-1)$ matrix. Let $B = \hat{Q}_{p-1}\hat{R}_{p-1}$. Then $A = [B\ a_p] = [\hat{Q}_{p-1}\ q_p] \begin{bmatrix} \hat{R}_{p-1} & \boldsymbol{r}_p \\ \mathbf{0} & r_{pp} \end{bmatrix}$, where q_p is an orthonormal vector to Q_{p-1}, $\boldsymbol{r}_p$ is a $(p-1)$ dimensional vector, is a valid QR factorization of A. Thus:

   $$A = [B\ a_p] = [\hat{Q}_{p-1}\hat{R}_{p-1} \quad \hat{Q}_{p-1}\boldsymbol{r}_p + r_{pp}q_p] = \hat{Q}\hat{R} \tag{1}$$

   Now let k diagonal entries of $\hat{R}$ are nonzero. There are two possible cases:

   i. Case 1: $r_{pp} \neq 0$. In this case only $k-1$ of the diagonal entries of $\hat{R}_{p-1}$ are nonzero. Using induction hypothesis, rank(B)$\geq k-1$. Also $a_p = \hat{Q}_{p-1}\boldsymbol{r}_p + r_{pp}q_p$ with $r_{pp} \neq 0$ and q_p does not lie in the span of columns of $\hat{Q}_{p-1}$. Therefore, a_p does not lie in the span of columns of $\hat{Q}_{p-1}$, which implies a_p does not lie in the span of columns of B. Hence, rank(A)=rank(B)+1 $\geq k$.

   ii. Case 2: $r_{pp} = 0$. In this case k of the diagonal entries of $\hat{R}_{p-1}$ are nonzero. Thus rank(B)$\geq k$. So trivialy, rank(A)$\geq k$.

5. Gram-Schmidt Process
   First we follow the steps of Algorithm 7.1. In the first iteration of the main loop we get that $r_{11} = fl(1+\epsilon^2) = 1$, and so $q_1 = v_1$ (the final value). In the second iteration we calculate $r_{12} = v_1^* v_2 = 1$ (exactly), and we get $(0, -\epsilon, \epsilon, 0)^*$ when updating v_2. The normalization step produces a very good approximation to $(0, -1/\sqrt{2}, 1/\sqrt{2}, 0)$, note that the fact that ϵ is small does not pose a problem. In the third iteration similarly to the above we calculate $r_{13} = 1$ and $r_{23} = 0$ exactly, and end up with a very good approximation of: $(0, -1/\sqrt{2}, 0, 1/\sqrt{2})$ for q_3. As we see q_2 and q_3 are clearly not close to be orthogonal, their inner product is near 0.5.

   Next, we follow the Modified Gram-Schmidt algorithm in which the calculation of r_{ij} is replaced with: $r_{ij} = q_i^* v_j$. In the first two iterations of the main loop the calculations are the same as above. In the third iteration however we still get $r_{13} = 1$, but $r_{23} = \epsilon/\sqrt{2}$ or actually a very good approximation of it in floating point arithmetic. (Note that the elements of q_2 were already approximations to $1/\sqrt{2}$ and $-1/\sqrt{2}$. The fact that r_{23} is small, poses no problem of forming q_3 as $(0, -\epsilon/\sqrt{6}, -\epsilon/\sqrt{6}, 0)$ and normalizing it to: $(0, -1/\sqrt{6}, -1/\sqrt{6}, 0)$. (Again these will be very good approximations but not the nearest floating point numbers necessarily, but that error is immaterial for this analysis.) Finally, we can verify that $q_2^* q_3$ is very close to 0 this time, while the other dot products are about $\epsilon/\sqrt{2}$ and $\epsilon/\sqrt{6}$.

   Note that q_1 and q_2 are not numerically orthogonal to each other as much as we would like them to be. For example when $\epsilon = 10^{-10}$ the dot product will be near $-7 \cdot 10^{-11}$.