MIT 18.06 Exam 2 Solutions, Spring 2017

Problem 1:

You are given the 6×6 matrix.

$$A = \begin{pmatrix} 1 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & -1 & 2 & -1 & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix}.$$

(a) Find the determinant of A.

Solution: since the determinant is not changed by elimination, we first do Gaussian elimination to put this in upper-triangular form. Similar to exam 1, elimination on this matrix follows a simple pattern:

$$\begin{pmatrix} \boxed{1} & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & -1 & 2 & -1 & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -1 & & & & \\ 0 & \boxed{1} & -1 & & & \\ & -1 & 2 & -1 & & \\ & & -1 & 2 & -1 & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix}.$$

$$\rightsquigarrow \cdots \rightsquigarrow \begin{pmatrix} 1 & -1 & & & & \\ 0 & 1 & -1 & & & \\ & 0 & 1 & -1 & & \\ & & 0 & 1 & -1 & \\ & & & 0 & 1 & -1 \\ & & & & 0 & 1 \end{pmatrix} = U$$

(Unlike in exam 1, our matrix A is nonsingular: the 2 in the lower-right corner of A gives a 1 in the lower-right corner of U.) No row swaps were required, and hence the determinant is just the product of the diagonals of U: $\boxed{\det A = 1}$.

(b) What is the projection matrix onto $C(A)$?

Solution: A is a nonsingular square matrix, so $\boxed{C(A) = \mathbb{R}^6}$ (the whole space).

(c) If you perform Gram–Schmidt orthogonalization on the columns of A, what is the pattern of nonzero entries in the resulting orthogonal matrix Q? (Note: this is *not* the same as Gram–Schmidt on U!)

Solution: in Gram–Schmidt, we subtract multiples of previous columns from subsequent ones. This will introduce new nonzero factors above the diagonal but not below the diagonal, leading to the nonzero pattern:

$$\boxed{Q = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ & \times & \times & \times & \times & \times \\ & & \times & \times & \times & \times \\ & & & \times & \times & \times \\ & & & & \times & \times \end{pmatrix}}.$$

That is, it is upper-triangular plus one additional nonzero entry below each diagonal. (This pattern is called an "upper Hessenberg" matrix.)

(If we want to be really careful here, we have to ensure that the Gram–Schmidt process does not introduce zeros above the diagonal by fortuitous cancellations. That doesn't happen for this matrix, but I don't expect you to verify this.)

Problem 2:

The equations of two lines in $\mathbb{R}^n$ are

$$\vec{y}_1(x_1) = \vec{a}_1 x_1 + \vec{b}_1$$
$$\vec{y}_2(x_2) = \vec{a}_2 x_2 + \vec{b}_2$$

where $\vec{a}_1, \vec{a}_2, \vec{b}_1, \vec{b}_2 \in \mathbb{R}^n$ and x_1 and x_2 are scalars. (On the exam, I originally wrote "$\vec{y}(x_1)$" and "$\vec{y}(x_2)$", which is notationally a bit ambiguous, but no one seemed to have been confused on that point.)

Write down a 2×2 system $C\vec{x} = \vec{d}$ of linear equations for $\vec{x} = (x_1, x_2)$ whose solution gives the (x_1, x_2) that **minimizes the distance between the two lines**. That is, find the entries of C and $\vec{d}$ (in terms of $\vec{a}_1, \vec{a}_2, \vec{b}_1, \vec{b}_2$) so that $\vec{x} = C^{-1}\vec{d}$ solves:

$$\min_{x_1, x_2} \|\vec{y}_1(x_1) - \vec{y}_2(x_2)\|.$$

Solution:

To start with, let's take the hint and write $\vec{y}_2(x_1) - \vec{y}_2(x_2)$ in terms of linear-algebra operations on $\vec{x}$:

$$\vec{y}_2(x_1) - \vec{y}_2(x_2) = \vec{a}_1 x_1 + \vec{b}_1 - \vec{a}_2 x_2 - \vec{b}_2 = \boxed{\underbrace{\begin{pmatrix} \vec{a}_1 & -\vec{a}_2 \end{pmatrix}}_{A} \vec{x} - \underbrace{(\vec{b}_2 - \vec{b}_1)}_{\vec{b}}},$$

where we have defined the $n \times 2$ matrix A whose columns are $\vec{a}_1$ and $-\vec{a}_2$ (note the sign!).

Now, we can see that we are really solving the least-square problem:

$$\min_{\vec{x} \in \mathbb{R}^2} \|A\vec{x} - \vec{b}\|$$

whose solution $\vec{x}$, from class, is given from the 2×2 normal equations:

$$\boxed{\underbrace{A^T A}_{C} \vec{x} = \underbrace{A^T \vec{b}}_{\vec{d}}}.$$

with

$$\boxed{\vec{d} = A^T(\vec{b}_2 - \vec{b}_1) = \begin{pmatrix} \vec{a}_1^T \\ -\vec{a}_2^T \end{pmatrix} (\vec{b}_2 - \vec{b}_1) = \begin{pmatrix} \vec{a}_1^T(\vec{b}_2 - \vec{b}_1) \\ -\vec{a}_2^T(\vec{b}_2 - \vec{b}_1) \end{pmatrix}}$$

and

$$\boxed{C = A^T A = \begin{pmatrix} \vec{a}_1^T \vec{a}_1 & -\vec{a}_1^T \vec{a}_2 \\ -\vec{a}_2^T \vec{a}_1 & \vec{a}_2^T \vec{a}_2 \end{pmatrix}}.$$

Common mistakes: Many people wrote something of the form $y_1 - y_2 = \begin{pmatrix} \vec{a}_1 & \\ & -\vec{a}_2 \end{pmatrix} \vec{x} - (\cdots)$, but this is wrong: $\begin{pmatrix} \vec{a}_1 & \\ & -\vec{a}_2 \end{pmatrix} \vec{x} = \begin{pmatrix} \vec{a}_1 x_1 \\ -\vec{a}_2 x_2 \end{pmatrix}$, and so you don't get $y_1 - y_2$ from this $2n \times 2$ matrix. Many people got the correct A and $\vec{b}$, but then wrote $A\vec{x} = \vec{b}$, which is a non-square problem that has no solution in general — you need to minimize $\|A\vec{x} - \vec{b}\|$ via least-squares. Many people got the correct A, but then wrote $A^T = \begin{pmatrix} \vec{a}_1 \\ -\vec{a}_2 \end{pmatrix}$ (which has the wrong size!) instead of $\begin{pmatrix} \vec{a}_1^T \\ -\vec{a}_2^T \end{pmatrix}$ — you then get terms like "$\vec{a}_1 \vec{a}_2$" and "$\vec{a}_1 \vec{b}_2$", and you should always be suspicious if your answer involves the nonsensical operation vector $\times$ vector!

Problem 3:

(a) If P projects onto $C(A^T)$, the row space of some $m \times n$ matrix A, then $(I - P)^2 x$ for any $x \in \mathbb{R}^n$ gives a vector in which fundamental subspace?

Solution: The matrix $I - P$ is the projection matrix onto the orthogonal complement $C(A^T)^\perp = N(A)$. The square of a projection matrix is the same matrix; explicitly: $(I-P)^2 = I^2 + P^2 - IP - PI = I + P - P - P = I - P$ since $P^2 = P$. Hence $\boxed{(I-P)^2 x \in N(A)}$.

(b) If A is a symmetric matrix and P is the projection matrix onto $N(A)$, what is PA?

Solution: since $A^T = A$, we have $N(A) = N(A^T) = C(A)^{\perp}$, and hence $\boxed{PA = 0}$ (the zero matrix). That is, every column of A is in $C(A)$, and hence P times each column (the result of PA) gives zero (the intersection of $C(A)$ and $N(A^T)$).

(c) If P is a permutation matrix, what is its QR factorization?

Solution: a permutation matrix already has orthonormal columns (they are just a permutation of the columns of I), and hence QR factorization does nothing. You simply get $\boxed{Q = P,\, R = I}$.

(d) If A and B are two matrices such that $A^T B = 0$ (the zero matrix), with QR factorizations $A = Q_A R_A$ and $B = Q_B R_B$, write down the QR factorization of the matrix $C = \begin{pmatrix} A & B \end{pmatrix}$ (that is, C is the columns of A followed by the colums of B) in terms of Q_A, Q_B, R_A, R_B.

Solution: $A^T B$ computes the dot products of all the columns of A (= rows of A^T) with all the columns of B. $A^T B = 0$ means that these are *all orthogonal*, i.e. $C(A) \perp C(B)$. Since Q_A and Q_B span the same space, $Q_A^T Q_B = 0$ as well. When we perform Gram–Schmidt on $C = \begin{pmatrix} A & B \end{pmatrix}$, we first orthonormalize the columns of A, getting Q_A again. Then we orthonormalize the columns of B, projecting out Q_A via $I - Q_A Q_A^T$. But since $\left(I - Q_A Q_A^T\right) B = B$ (the columns of B are already orthogonal to the A columns), the projection with Q_A does nothing. We will end up just doing ordinary Gram-Schmidt on B, obtaining Q_B as before, to finally obtain the QR factorization

$$\boxed{C = \underbrace{\begin{pmatrix} Q_A & Q_B \end{pmatrix}}_{Q} \underbrace{\begin{pmatrix} R_A & \\ & R_B \end{pmatrix}}_{R}} = \begin{pmatrix} Q_A R_A & Q_B R_B \end{pmatrix} = \begin{pmatrix} A & B \end{pmatrix}.$$

It is easy to check that this is indeed a QR factorization of C. R is upper-triangular and $Q^T Q = \begin{pmatrix} Q_A^T \\ Q_B^T \end{pmatrix} \begin{pmatrix} Q_A & Q_B \end{pmatrix} = \begin{pmatrix} Q_A^T Q_A & Q_A^T Q_B \\ Q_B^T Q_A & Q_B^T Q_B \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} =$ identity, as desired.