18.06 - Recitation 6 - Solutions

Sam Turton

April 2, 2019

1 Review problems for midterm 2

Problem 1.
The matrix A has a nullspace $N(A)$ spanned by

$$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

and a left nullspace $N(A^T)$ spanned by

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}.$$

(a) What is the **shape** of the matrix A and what is its **rank**?

(b) If we consider the vector

$$b = \begin{pmatrix} -1 \\ \alpha \\ 0 \\ \beta \end{pmatrix},$$

for **what value(s)** of α and β (if any) is $Ax = b$ solvable? Will the solution (if any) be **unique**?

(c) Give the orthogonal **projections** of

$$y = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$$

onto **two** of the four fundamental subspaces of A.

Solution

(a) Since $N(A)$ is a subspace of $\mathbb{R}^3$, the matrix A must have three columns. Since $N(A^T)$ is a subspace of $\mathbb{R}^4$, the matrix A must have four rows. So A is a 4×3 matrix. The matrix has 3 columns and the null space has dimension 1, and so the rank of the matrix is $r = 3 - 1 = 2$.

(b) If $Ax = b$ is solvable, then $b \in C(A)$. Since $C(A)$ is the orthogonal complement of $N(A^T)$, this means that an equivalent condition for $Ax = b$ to be solvable is that b is orthogonal to $N(A^T)$. This gives us two constraints on b:

$$b^T \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = 0 \quad \Longrightarrow \quad -1 + \alpha + \beta = 0,$$

$$b^T \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} = 0 \quad \Longrightarrow \quad -1 + \alpha - \beta = 0.$$

And so $b \in C(A)$ requires $\alpha = 1,\ \beta = 0$. For these values of $\alpha$and β, the solution of $Ax = b$ is not unique, since $N(A)$ has dimension 1. Given any particuar solution of $Ax = b$, we can add on any multiple of $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ and the resulting vector would still be a solution.

(c) The vector $y = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$ is in $\mathbb{R}^3$, and so we can project onto $N(A)$ and $C(A^T)$. To project onto $N(A)$, we use the formula to project y onto $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$:

$$p_{N(A)} = \frac{\begin{pmatrix} 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}}{\begin{pmatrix} 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}.$$

To compute the projection onto $C(A^T)$, recall that if $p = Py$ is the projection of y onto some subspace, then $(I - P)y$ will project y onto the orthogonal complement of this subspace. Since $C(A^T)$ is orthogonal to $N(A)$, the projection of y onto $C(A^T)$ is given by:

$$\begin{aligned} p_{C(A^T)} &= \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} - p_{N(A)} \\ &= \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} \\ &= \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}. \end{aligned}$$

Problem 2.
You have a matrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

(a) Give the **ranks** of A, A^T, and A^TA, and also give **bases** for $C(A)$, $N(A)$, and $N(A^TA)$. (Look carefully at the columns of A, since very little calculation is needed!)

(b) Suppose we are looking for a least squares solution $\hat{x}$ that minimizes $\|b - Ax\|$ for $b = \begin{pmatrix} 0 \\ 2 \\ 1 \\ -1 \end{pmatrix}$. At this minimum, $p = A\hat{x}$ will be the projection of b onto .... ? **Find** p.

Solution

(a) The first and third columns of A are the same, while the first and second columns are linearly independent. This means that the rank of A is 2. The rank of $A^T$and the rank of A^TA are equal to the rank of A. A basis for $C(A)$ is then just the first two columns $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix}$. The nullspace of A is one dimensional, and since the first and third columns are the same, a basis for $N(A)$ is given by the vector $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Finally, $N(A) = N(A^TA)$, and so our basis for $N(A)$ is also a basis for $N(A^TA)$.

(b) Firstly, $p = A\hat{x}$ is the projection of b onto $C(A)$. To find $\hat{x}$ we must solve the normal equations $A^TA\hat{x} = A^Tb$. However, since A only has two linearly independent columns we can simplify our calculations by instead using the matrix $B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}$, and solve the normal equations $B^TB\hat{x} = B^Tb$. We can calculate

$$B^TB = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 3 & 6 \end{pmatrix}$$

and

$$B^Tb = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}.$$

The normal equations are then:

$$\begin{pmatrix} 3 & 3 \\ 3 & 6 \end{pmatrix} \hat{x} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$$
$$\implies \hat{x} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Finally, we can compute

$$p = B\hat{x} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix}.$$

Problem 3.

(a) Show that the trace of $A^T A$ must always be ≥ 0 by deriving a simple formula for trace$(A^T A)$ in terms of the matrix entries a_{ij} (i-th row, j-th column) of A. This is called the *Frobenius norm*

$$\|A\|_F = \sqrt{\text{trace}(A^T A)}$$

of the matrix.

(b) Using the compact SVD $A = U\Sigma V^T$, derive a simple relationship between the Frobenius norm $\|A\|_F$ and the singular values $\sigma_1, \ldots, \sigma_r$ of A.

Solution

(a) Suppose A is an $m \times n$ matrix. The trace of $A^T A$ is the sum of the n diagonal entries $(A^T A)_{ii}$. Each of these diagonal entries is given by the sum $(A^T A)_{ii} = \sum_{j=1}^{m} a_{ji}^2$. So

$$\boxed{\text{trace}(A^T A) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ji}^2}, \tag{1}$$

which is necessarily ≥ 0 since every term in this sum is squared.

(b) We can use the SVD to derive a simple relationship between the Frobenius norm $\|A\|_F$ and the singular values by considering trace$(A^T A)$:

$$\begin{align}
\text{trace}(A^T A) &= \text{trace}\left[(U\Sigma V^T)^T (U\Sigma V^T)\right] \tag{2} \\
&= \text{trace}\left[V\Sigma^T U^T U\Sigma V^T\right] \tag{3} \\
&= \text{trace}\left[V\Sigma\Sigma V^T\right] \tag{4} \\
&= \text{trace}\left[\Sigma V^T V\Sigma\right] \tag{5} \\
&= \text{trace}\left[\Sigma^2\right] \tag{6} \\
&= \sum_{i=1}^{r} \sigma_i^2 \tag{7}
\end{align}$$

And so

$$\boxed{\|A\|_F = \sqrt{\sum_{i=1}^{r} \sigma_i^2}} \tag{8}$$

Problem 4.

(a) If Q is an orthogonal matrix ($Q^T = Q^{-1}$), explain why it follows from the rules for determinants that $\det Q$ must be .... or ....?

(b) If P is a 3×3 projection matrix onto a 2d subspace, then its determinant must be ....?

(c) An anti-symmetric matrix is a $n \times n$ matrix A with $A^T = -A$. What is $\det A$ when n is odd?

Solution

(a) If Q is an orthogonal, square matrix then we know that $Q^T = Q^{-1}$. By the rules of determinants, we know that $\det Q^T = \det Q$ and that $\det Q^{-1} = \frac{1}{\det Q}$. We can then equate these two expressions, to deduce that $\det Q = \frac{1}{\det Q} \implies (\det Q)^2 = 1$, which means that $\det Q = \pm 1$.

(b) If P is a 3×3 projection matrix onto a 2d subspace, then P has rank 2. This means that one of the pivots of P will be zero, and so $\det P = 0$.

(c) If A is an $n \times n$ matrix, then $\det(-A) = (-1)^n \det A$. If A is skew symmetric, then $A^T = -A \implies \det A^T = \det(-A) \implies \det A = (-1)^n \det A$. If m is odd, then this necessarily means $\det A = 0$. However, if n is even, then generally $\det A \neq 0$.