CAAM 335 Matrix Analysis

Solutions to Homework 10

Problem 1 (20 points) Consider

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

(a) (6 points)
Compute all the eigenvalues and orthonormal eigenvectors of AA^T and A^TA (many quantities can be easily obtained without substantial calculations).

(b) (7 points)
Find the singular value decomposition $A = U\Sigma V^T$ by hand and show your steps (you may compare with the result from the MATLAB command `svd`).

(c) (7 points)
Without any further computation, give an orthonormal basis for each of the subspaces: $\mathcal{R}(A^T)$, $\mathcal{N}(A)$, $\mathcal{R}(A)$, $\mathcal{N}(A^T)$, if it is non-trivial.

Solution:

(a)

$$AA^T = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix}$$

Since AA^T is diagonal, the eigenvalues are 2 and 1 with corresponding eigenvectors $(1,0)^T$ and $(0,1)^T$.

$$A^TA = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

As one would expect, the rank of A^TA is 2, columns 1 and 3 are linearly dependent, thus 0 is an eigenvalue with eigenvector $(1,0,1)^T$. One can also see that 1 is an eigenvalue with eigenvector $(0,1,0)^T$. Since A^TA is a symmetric matrix and the eigenvectors of symmetric matrix are orthogonal, the only remaining direction in $\mathbb{R}^3$ must also be an eigenvector, thus $(1,0,-1)^T$ is an eigenvector with eigenvalue 2.

Another way of getting eigenvectors of A^TA corresponding to nonzero eigenvalues is to multiply the eigenvectors of AA^T, $(1,0)^T$ and $(0,1)^T$, by A^T. In this case, the two columns of A^T turn out to be eigenvectors of A^TA corresponding to $\lambda_1 = 2$ and $\lambda_2 = 1$.

(b) From $A = U\Sigma V^T$, the eigenvectors of A^TA give the right singular vectors and the eigenvectors of AA^T give the left singular vectors. The singular values involve taking the square root of eigenvalues of either of AA^T. Hence,

$$V = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

(c) From the properties of the SVD the range of A is the span of the columns of U associated with the nonzero eigenvalues of AA^T. Similarly the range of A^T is equal to the span of the columns of V corresponding to nonzero eigenvalues of A^TA. The null space of A is the span of the columns of U associated with the zero eigenvalues of AA^T. Similarly, the null space of A^T is the span of the columns of V associated with the zero eigenvalues of A^TA.

$$\mathcal{R}(A^T) = \text{span}\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}, \qquad \mathcal{N}(A) = \text{span}\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\},$$

$$\mathcal{R}(A) = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \qquad \mathcal{N}(A^T) = \{\mathbb{0}\},$$

where $\mathcal{N}(A^T)$ technically is not empty, rather it contains only the zero vector which is contained in every subspace.

Problem 2 (10 points) The pseudo-inverse of A is defined as $A^\dagger = V\Sigma^\dagger U^T$. Is it true or false that $AA^\dagger A = A$ and $A^\dagger AA^\dagger = A^\dagger$? Justify your answers.

Solution:

The statement is true.
Let $A \in \mathbb{R}^{m\times n}$ with rank $r \leq \min(n,m)$. Then $A = U\Sigma V^T$ with $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ and $\Sigma \in \mathbb{R}^{m\times n}$.

$$AA^\dagger A = U\Sigma V^T V\Sigma^\dagger U^T U\Sigma V^T = U\Sigma\Sigma^\dagger\Sigma V^T$$

$$\Sigma\Sigma^\dagger\Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \mathbf{0}_{r\times(n-r)} \\ & & \sigma_r & \\ & \mathbf{0}_{(m-r)\times r} & & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma_1} & & & \\ & \ddots & & \mathbf{0}_{r\times(m-r)} \\ & & \frac{1}{\sigma_r} & \\ & \mathbf{0}_{(n-r)\times r} & & \mathbf{0}_{(n-r)\times(m-r)} \end{pmatrix} \Sigma$$

$$= \begin{pmatrix} I_r & \\ & 0_{(m-r)\times(m-r)} \end{pmatrix} \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \mathbf{0}_{r\times(n-r)} \\ & & \sigma_r & \\ & \mathbf{0}_{(m-r)\times r} & & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \mathbf{0}_{r\times(n-r)} \\ & & \sigma_r & \\ & \mathbf{0}_{(m-r)\times r} & & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix} = \Sigma$$

Hence

$$AA^\dagger A = U\Sigma V^T = A.$$

$$A^\dagger AA^\dagger = V\Sigma^\dagger U^T U\Sigma V^T V\Sigma^\dagger U^T = V\Sigma^\dagger\Sigma\Sigma^\dagger U^T$$

Using the result above for $\Sigma\Sigma^\dagger$,

$$\Sigma^\dagger\Sigma\Sigma^\dagger = \begin{pmatrix} \frac{1}{\sigma_1} & & & \\ & \ddots & & \mathbf{0}_{r\times(m-r)} \\ & & \frac{1}{\sigma_r} & \\ & \mathbf{0}_{(n-r)\times r} & & \mathbf{0}_{(n-r)\times(m-r)} \end{pmatrix} \begin{pmatrix} I_r & \\ & 0_{(m-r)\times(m-r)} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{\sigma_1} & & & \\ & \ddots & & \mathbf{0}_{r\times(m-r)} \\ & & \frac{1}{\sigma_r} & \\ & \mathbf{0}_{(n-r)\times r} & & \mathbf{0}_{(n-r)\times(m-r)} \end{pmatrix} = \Sigma^\dagger$$

Thus

$$A^\dagger AA^\dagger = V\Sigma^\dagger U^T = A^\dagger$$

Problem 3 (10 points) Let matrix $A \in \mathbb{R}^{m\times n}$. Prove that both $AA^\dagger$ and $A^\dagger A$ are orthogonal projection matrices (idempotent and symmetric). Which subspaces do they project onto, respectively? Justify your answers.

Solution:

To be an orthogonal projector a matrix must be both idempotent, $P^2 = P$, and symmetric $P^T = P$.
For the idempotent property,

$$(AA^\dagger)(AA^\dagger) = AA^\dagger AA^\dagger = A(A^\dagger AA^\dagger) = AA^\dagger,$$

where we have utilized the result from the previous exercise and similarly

$$(A^\dagger A)(A^\dagger A) = A^\dagger AA^\dagger A = (A^\dagger AA^\dagger)A = A^\dagger A.$$

For symmetry,

$$AA^\dagger = U\Sigma V^T V\Sigma^\dagger U^T = U\Sigma\Sigma^\dagger U^T = U\begin{pmatrix} I_r & \\ & \mathbf{0}_{m-r} \end{pmatrix} U^T,$$

which is clearly symmetric and similarly

$$A^\dagger A = V\Sigma^\dagger U^T U\Sigma V^T = V\Sigma^\dagger\Sigma V^T = V\begin{pmatrix} I_r & \\ & \mathbf{0}_{n-r} \end{pmatrix} V^T.$$

From the decompositions above for $AA^\dagger$ and $A^\dagger A$, one can see that $AA^\dagger$ projects onto the range of A whereas $A^\dagger A$ projects onto the range of A^T.

Problem 4 (10 points) Let matrix $A \in \mathbb{R}^{m\times n}$ have rank n. Show that in this case $A^\dagger = (A^T A)^{-1} A^T$ and $A^\dagger A = I$ (so $A^\dagger$ is also called left inverse). Give the analogous results for the case when A has rank m.

Solution:

$$(A^T A)^{-1} A^T = (V\Sigma^T U^T U\Sigma V)^{-1} V\Sigma^T U^T = V(\Sigma^T\Sigma)^{-1} V^T V\Sigma^T U^T = V(\Sigma^T\Sigma)^{-1}\Sigma^T U^T$$

Since the rank of A is n, and $\Sigma^T\Sigma$ is invertible, thus

$$\begin{aligned}
(\Sigma^T\Sigma)^{-1}\Sigma^T &= \left(\begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \mathbf{0}_{n\times(m-n)} \\ & & \sigma_n & \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & \mathbf{0}_{(m-n)\times n} & \end{pmatrix} \right)^{-1} \Sigma^T \\
&= \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{pmatrix}^{-1} \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \mathbf{0}_{n\times(m-n)} \\ & & \sigma_n & \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{\sigma_1} & & & \\ & \ddots & & \mathbf{0}_{n\times(m-n)} \\ & & \frac{1}{\sigma_n} & \end{pmatrix} = \Sigma^\dagger,
\end{aligned}$$

thus

$$(A^TA)^{-1}A^T = V\Sigma^\dagger U^T = A^\dagger.$$

Since (A^TA) is invertible because the rank of A is n,

$$A^\dagger A = (A^TA)^{-1}A^TA = (A^TA)^{-1}(A^TA) = I.$$

For the case when A has rank m, $A^\dagger = A^T(AA^T)^{-1}$. Similar to the calculations above,

$$A^T(AA^T)^{-1} = V\Sigma^TU^T(U\Sigma V^TV\Sigma^TU^T)^{-1} = V\Sigma^TU^TU(\Sigma\Sigma^T)^{-1}U^T = V\Sigma^T(\Sigma\Sigma^T)^{-1}U^T$$

Since the rank of A is m, $\Sigma\Sigma^T$ is invertible

$$\begin{aligned}
\Sigma^T(\Sigma\Sigma^T)^{-1} &= \Sigma^T\left(\begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \mathbf{0}_{m\times(n-m)} \\ & & \sigma_m & \end{pmatrix}\begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \\ & \mathbf{0}_{(n-m)\times m} & \end{pmatrix}\right)^{-1} \\
&= \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & \mathbf{0}_{(n-m)\times m} & \end{pmatrix}\begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{pmatrix}^{-1} \\
&= \begin{pmatrix} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_m} \\ & \mathbf{0}_{(n-m)\times m} & \end{pmatrix} = \Sigma^\dagger,
\end{aligned}$$

thus

$$A^T(AA^T)^{-1} = V\Sigma^\dagger U^T = A^\dagger.$$

Since (A^TA) is invertible because the rank of A is m,

$$AA^\dagger = AA^T(AA^T)^{-1} = (AA^T)(AA^T)^{-1} = I.$$