E2 212: Matrix Theory (Fall 2010)
Solutions to Test - 1

1. Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n] \in \mathbb{R}^{m \times n}$ be a tall matrix. Let $\mathcal{S} \triangleq \mathcal{R}(\mathbf{X})$, and let P be an orthogonal projector onto $\mathcal{S}$.

   (a) If $\mathbf{X}$ is full rank, show that P can be written as $P = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$. What is $\|P\|_2$? (4 points)

   (b) Suitably modify P when $\mathbf{X}$ is not full rank. *Hint:* Let Q be an orthonormal basis for $\mathcal{R}(\mathbf{X})$. (4 points)

   (c) Let $\mathbf{X}$ be full rank and let B be a square matrix such that $B^T B = \mathbf{X}^T\mathbf{X}$. What special property does the matrix $\mathbf{X}B^{-1}$ possess? Using this, what can you say about the matrix $P = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$? (4 points)

   (d) Show that the rank of P is equal to its trace. (4 points)

   (e) What are the singular values of P? (4 points)

Solution

(a) Recall that by definition, a matrix $P \in \mathbb{R}^{m \times m}$ is a projector onto $\mathcal{S}$ if (a) $\mathcal{R}(P) = \mathcal{S}$, (b) $P^2 = P$ and (c) $P^T = P$. Since $\mathbf{X}$ has full column rank, $(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$ spans $\mathbb{R}^m$, and hence, $\mathcal{R}(P) = \mathcal{S}$. The other two properties are trivial to verify. Also, $\|P\|_2 = 1$ because its eigenvalues are all 1 or 0, so the largest eigenvalue is 1.

(b) It is simple to verify that QQ^T is the required projection. It satisfies the three properties required. Alternatively, if X_{sub} is the matrix constructed from the largest possible subset of linearly independent columns of X (think Gram-Schmidt procedure), then the matrix $P = X_{sub}(X_{sub}^T X_{sub})^{-1} X_{sub}^T$ is the required projector.

(c) $(\mathbf{X}B^{-1})^T\mathbf{X}B^{-1} = I$, hence $\mathbf{X}B^{-1}$ contains orthonormal columns. Moreover, its columns span $\mathbf{X}$ since B^{-1} is full rank. By (b), this implies that $P =$

$\mathbf{X}B^{-1}(\mathbf{X}B^{-1})^T = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ is a projector onto $\mathcal{R}(X)$.

(d) There are multiple ways of solving this. One way is to write $P = QQ^T$ where Q has orthonormal columns. Then, rank(P) = the number of linearly independent columns = trace(P), since the latter equals trace(Q^TQ) = trace(I_r), where r is the rank (and the number of columns) of Q.

(e) The singular values are 1 and 0. (They are the same as the eigenvalues since $P^2 = P$.)

2. Let A and B be $n \times n$ matrices that commute: $AB = BA$. If λ is an eigenvalue of A, let $\mathbf{V}_\lambda$ be the subspace of all eigenvectors corresponding to this eigenvalue.

   (a) Prove that the eigenspace $\mathbf{V}_\lambda$ is an invariant subspace of the matrix B, i.e., show that $\forall\, \mathbf{v} \in \mathbf{V}_\lambda$, $B\mathbf{v} \in \mathbf{V}_\lambda$. (4 points)

   (b) Show that $\exists\, \mathbf{v} \in \mathbf{V}_\lambda$, such that $\mathbf{v}$ is an eigenvector of B. (12 points)

   Solution

   (a) Let $\mathbf{v}$ be an eigenvector of A, i.e., $BA\mathbf{v} = AB\mathbf{v} = \lambda B\mathbf{v}$, where λ is an eigenvalue of A. Therefore, $B\mathbf{v}$ is an eigenvector of A. Thus, $\forall\, \mathbf{v} \in \mathbf{V}_\lambda$, $B\mathbf{v} \in \mathbf{V}_\lambda$.

   (b) See "Matrix Analysis" by Horn and Johnson, page no. 51, lemma 1.3.17.

3. Let A be a $n \times n$ matrix and $\mathbf{z}$ be a vector with the property that $A^{k-1}\mathbf{z} \neq 0$ but $A^k\mathbf{z} = 0$. Show that $\mathbf{z}, A\mathbf{z}, \ldots, A^{k-1}\mathbf{z}$ are linearly independent. (8 points)

   Solution
   We need to prove that $\alpha_1\mathbf{z} + \alpha_2 A\mathbf{z}^2 + \ldots + \alpha_k A^{k-1}\mathbf{z} = 0 \Rightarrow \alpha_i = 0,\ i = 1, \ldots, k$. Consider, $\alpha_1\mathbf{z} + \alpha_2 A\mathbf{z}^2 + \ldots + \alpha_k A^{k-1}\mathbf{z} = 0$. Applying A^{k-1} on it, we get

$$A^{k-1}\left[\alpha_1\mathbf{z} + \alpha_2 A\mathbf{z} + \ldots + \alpha_{k-1}A^{k-2}\mathbf{z}\right] = 0 \Rightarrow \alpha_1 A^{k-1}\mathbf{z} = 0 \Rightarrow \alpha_1 = 0.$$

   Setting $\alpha_1 = 0$ and multiplying the equation by A^{k-2}, we can show that $\alpha_2 = 0$. Continuing this way, we get, $\alpha_1 = \alpha_2 = \ldots \alpha_{k-1} = 0$.

4. There are no square matrices A, B with the property that $AB - BA = I$. Prove or give counterexample. (8 points)

   Solution
   Let us assume such matrices of size $N \times N$ exist. Taking trace on both sides and using

the fact Trace(AB) = Trace(BA), we get, 0 on the right hand side and N on the left hand side, which is contradiction.

5. What is wrong with this argument for the proof of the Cayley-Hamilton theorem? "Since $p_A(t) = \det(tI - A)$, $p_A(A) = \det(AI - A) = 0$. Therefore, $p_A(A) = 0$". (4 points)

Solution
One has to write out the polynomial and then substitute for t.

6. Consider the Householder reflection

$$H_{\mathbf{x}} = I - 2\frac{\mathbf{x}\mathbf{x}^T}{\mathbf{x}^T\mathbf{x}},$$

$\mathbf{x} \in \mathbb{R}^n$.

(a) Show that $H_{\mathbf{x}}\mathbf{x} = -\mathbf{x}$ and if $\mathbf{y} \in \mathbb{R}^{n\times 1}$ is such that $\mathbf{y}^T\mathbf{x} = \mathbf{x}^T\mathbf{y} = 0$, then $H_{\mathbf{x}}\mathbf{y} = \mathbf{y}$. (4 points)

(b) Show that $H_{\mathbf{x}}$ is symmetric and orthogonal. What is the trace of $H_{\mathbf{x}}$? (4 points)

(c) Given linearly independent vectors $\mathbf{y}, \mathbf{z} \in \mathbb{R}^{n\times 1}$, find $\mathbf{x}$ as a linear combination of $\mathbf{y}$ and $\mathbf{z}$ such that $H_{\mathbf{x}}\mathbf{y} \in \text{span}(\mathbf{z})$. (8 points)

Solution

(a)

$$H_{\mathbf{x}}\mathbf{x} = \mathbf{x} - 2\frac{\mathbf{x}(\mathbf{x}^T\mathbf{x})}{\mathbf{x}^T\mathbf{x}} = -\mathbf{x}$$

(b)

$$\begin{aligned}
H_{\mathbf{x}}^T &= (I - 2\frac{\mathbf{x}\mathbf{x}^T}{\mathbf{x}^T\mathbf{x}})^T = I - 2\frac{(\mathbf{x}\mathbf{x}^T)^T}{\mathbf{x}^T\mathbf{x}} \\
&= I - 2\frac{\mathbf{x}\mathbf{x}^T}{\mathbf{x}^T\mathbf{x}} = H_{\mathbf{x}} \\
H_{\mathbf{x}}^T H_{\mathbf{x}} &= (I - 2\frac{\mathbf{x}\mathbf{x}^T}{\mathbf{x}^T\mathbf{x}})^T(I - 2\frac{\mathbf{x}\mathbf{x}^T}{\mathbf{x}^T\mathbf{x}}) \\
&= I - 2\frac{\mathbf{x}\mathbf{x}^T}{\mathbf{x}^T\mathbf{x}} - 2\frac{\mathbf{x}\mathbf{x}^T}{\mathbf{x}^T\mathbf{x}} + 4\frac{(\mathbf{x}\mathbf{x}^T)(\mathbf{x}\mathbf{x}^T)}{(\mathbf{x}^T\mathbf{x})^2} \\
&= I
\end{aligned}$$

Also, $\text{tr}(H_{\mathbf{x}}) = n-2$ is easy to show since the eigenvalues are $+1$ (with multiplicity $n-1$) and -1 (with multiplicity 1), and the trace equals the sum of the eigenvalues.

(c) Let $\mathbf{x} = a\mathbf{y} + b\mathbf{z}$

$$
\begin{aligned}
H_{\mathbf{x}}\mathbf{y} &= \mathbf{y} - 2\frac{(a\mathbf{y}+b\mathbf{z})(a\mathbf{y}+b\mathbf{z})^T}{(a\mathbf{y}+b\mathbf{z})^T(a\mathbf{y}+b\mathbf{z})}\mathbf{y} \\
&= \mathbf{y} - 2\frac{a^2(\mathbf{y}^T\mathbf{y})\mathbf{y} + ab(\mathbf{z}^T\mathbf{y})\mathbf{z} + ab(\mathbf{z}^T\mathbf{y})\mathbf{y} + b^2(\mathbf{z}^T\mathbf{y})\mathbf{z}}{a^2\mathbf{y}^T\mathbf{y} + 2ab\mathbf{y}^T\mathbf{z} + b^2\mathbf{z}^T\mathbf{z}} \\
&= \left(1 - 2\frac{a^2(\mathbf{y}^T\mathbf{y}) + ab(\mathbf{z}^T\mathbf{y})}{a^2\mathbf{y}^T\mathbf{y} + 2ab\mathbf{y}^T\mathbf{z} + b^2\mathbf{z}^T\mathbf{z}}\right)\mathbf{y} - 2\left(\frac{ab(\mathbf{z}^T\mathbf{y})\mathbf{z} + b^2(\mathbf{z}^T\mathbf{y})}{a^2\mathbf{y}^T\mathbf{y} + 2ab\mathbf{y}^T\mathbf{z} + b^2\mathbf{z}^T}\right)\mathbf{z}
\end{aligned}
$$

For this vector to lie in span($\mathbf{z}$), the coefficient of $\mathbf{y}$ must be zero. Thus we get,

$$-a^2\mathbf{y}^T\mathbf{y} + b^2\mathbf{z}^T\mathbf{z} = 0$$

One of the possible solution is $a = 1$ and $b = \frac{\|\mathbf{y}\|_2}{\|\mathbf{z}\|_2}$.

7. Let $A \in \mathbb{R}^{m\times n}$ with $m \geq n$, $\mathbf{y} \in \mathbb{R}^n$, define

$$\tilde{A} = \begin{bmatrix} A \\ \mathbf{y}^T \end{bmatrix}.$$

Show that $\sigma_n(\tilde{A}) \geq \sigma_n(A)$ and $\sigma_1(\tilde{A}) \leq \sqrt{\|A\|_2^2 + \|\mathbf{y}\|_2^2}$. (12 points)

Solution
First, note that $\tilde{A}^T\tilde{A} = A^TA + \mathbf{y}\mathbf{y}^T$. Also, $\sigma_n^2(A) = \lambda_{\min}(A^TA)$, $\sigma_n^2(\tilde{A}) = \lambda_{\min}(\tilde{A}^T\tilde{A})$. Now,

$$
\begin{aligned}
\lambda_{\min}(\tilde{A}^T\tilde{A}) &= \min_{\|\mathbf{x}\|=1} \mathbf{x}^T\tilde{A}^T\tilde{A}\mathbf{x} \\
&= \min_{\|\mathbf{x}\|=1} \mathbf{x}^TA^TA\mathbf{x} + \left|\mathbf{x}^T\mathbf{y}\right|^2 \\
&\geq \min_{\|\mathbf{x}\|=1} \mathbf{x}^TA^TA\mathbf{x} = \lambda_{\min}(A^TA),
\end{aligned}
$$

where the last inequality is because the term $\left|\mathbf{x}^T\mathbf{y}\right|$ is always non-negative. Thus, $\sigma_n(\tilde{A}) \geq \sigma_n(A)$. To show the second part,

$$
\begin{aligned}
\sigma_1^2(\tilde{A}) &= \max_{\|\mathbf{x}\|=1}\left[\mathbf{x}^TA^TA\mathbf{x} + \left|\mathbf{x}^T\mathbf{y}\right|^2\right] \\
&\leq \left[\max_{\|\mathbf{x}\|=1} \mathbf{x}^TA^TA\mathbf{x}\right] + \left[\max_{\|\mathbf{x}\|=1}\left|\mathbf{x}^T\mathbf{y}\right|^2\right] = \|A\|_2^2 + \|\mathbf{y}\|_2^2,
\end{aligned}
$$

Thus, $\sigma_1^2(\tilde{A}) \leq \|A\|_2^2 + \|\mathbf{y}\|_2^2$, which proves the required result.

8. (a) Show that any square matrix which commutes with a diagonal matrix, whose all diagonal entries are different, is diagonal. (4 points)

   (b) Two matrices are called simultaneously diagonalizable, if they are diagonalized by the same matrix. Show that simultaneously diagonalizable matrices commute. (4 points)

(c) If two matrices commute and one of them is similar to a diagonal matrix with all the diagonal entries different, show that they are simultaneously diagonalizable. (8 points)

Solution

(a) Let $A = [a_{ij}]$ be a square matrix which commutes with diagonal matrix $D = \text{diag}(d_1, \ldots, d_n)$, i.e, $AD = DA$. Consider (i,j)th element of left hand side and right hand side.

$$\begin{aligned} \sum_{l=1}^{n} a_{il} d_{lj} &= \sum_{l=1}^{n} d_{il} a_{lj} \\ \implies a_{ij} d_{jj} &= d_{ii} a_{ij} \end{aligned}$$

Since all the diagonal entries of D are different, it follows that $a_{ij} = 0$ for $i \neq j$. Hence A is a diagonal matrix.

(b) Let A and B be two simultaneously diagonalizable matrices. Let $SAS^{-1} = D_1$ and $SBS^{-1} = D_2$, D_1 and D_2 diagonal matrices. Since diagonal matrices always commute,

$$\begin{aligned} AB &= S^{-1}D_1SS^{-1}D_2S = S^{-1}D_1D_2S \\ &= S^{-1}D_2D_1S = S^{-1}D_2SS^{-1}D_1S \\ &= BA \end{aligned}$$

(c) Let $AB = BA$ and $SAS^{-1} = D_1$, D_1 a diagonal matrix.

$$\begin{aligned} AB &= BA \\ \implies S^{-1}D_1SB &= BS^{-1}D_1S \\ \implies D_1SBS^{-1} &= SBS^{-1}D_1 \end{aligned}$$

SBS^{-1} commutes with a diagonal matrix with all diagonal entries different. Thus SBS^{-1} is also diagonal.