MATH 110: LINEAR ALGEBRA
SPRING 2007/08
PROBLEM SET 9 SOLUTIONS

1. A matrix $S \in \mathbb{R}^{n\times n}$ is called *skew symmetric* if $S^\top = -S$.

(a) For any matrix $A \in \mathbb{R}^{n\times n}$ for which $I + A$ is nonsingular, show that
$$(I - A)(I + A)^{-1} = (I + A)^{-1}(I - A). \tag{1.1}$$
We will write
$$\frac{I - A}{I + A}$$
for the matrix in (1.1). [Note: In general, $AB^{-1} \neq B^{-1}A$ and so
$$\frac{A}{B}$$
is ambiguous since it could mean either AB^{-1} or $B^{-1}A$.]

Solution. Note that
$$(I - A)(I + A)^{-1} = (I + A)^{-1}(I - A)$$
iff
$$(I + A)(I - A)(I + A)^{-1} = (I - A)$$
iff
$$(I + A)(I - A) = (I - A)(I + A)$$
and this last equation is evidently true since both sides equal $I - A^2$.

(b) Let $Q \in \mathbb{R}^{n\times n}$ be an orthorgonal matrix such that $I + Q$ is nonsingular. Show that
$$\frac{I - Q}{I + Q}$$
is a skew symmetric matrix.

Solution. Let $S := (I + Q)^{-1}(I - Q)$. Since $(A^{-1})^\top = (A^\top)^{-1}$ for any nonsingular matrix A, and since $Q^\top Q = I = QQ^\top$, we get
$$\begin{aligned} S^\top &= (I - Q)^\top[(I + Q)^{-1}]^\top \\ &= (I - Q^\top)(I + Q^\top)^{-1} \\ &= (QQ^\top - Q^\top)(QQ^\top + Q^\top)^{-1} \\ &= [(Q - I)Q^\top][(Q + I)Q^\top]^{-1} \\ &= (Q - I)Q^\top(Q^\top)^{-1}(Q + I)^{-1} \\ &= (Q - I)(Q + I)^{-1} \\ &= -S. \end{aligned}$$
So S is skew symmetric.

(c) Let $S \in \mathbb{R}^{n\times n}$ be a skew symmetric matrix. Show that
$$\frac{I - S}{I + S}$$
is an orthogonal matrix.

Date: May 16, 2008 (Version 1.0).

SOLUTION. Let $Q := (I+S)^{-1}(I-S)$. Since $(A^{-1})^\top = (A^\top)^{-1}$ for any nonsingular matrix A, and since $S^\top = -S$, we get,

$$\begin{aligned} Q^\top &= (I-S)^\top[(I+S)^{-1}]^\top \\ &= (I-S^\top)(I+S^\top)^{-1} \\ &= (I+S)(I-S)^{-1} \\ &= [(I-S)(I+S)^{-1}]^{-1} \\ &= Q^{-1}. \end{aligned}$$

(d) Why is it unnecessary to require that $I+S$ be nonsingular in (c)? [Hint: Problem **3** below.]
SOLUTION. By Problem **3**(d),

$$\mathbf{x}^\top(I+S)\mathbf{x} = \mathbf{x}^\top I\mathbf{x}$$

for all $\mathbf{x} \in \mathbb{R}^n$. Since $\mathbf{x}^\top I\mathbf{x} = \mathbf{x}^\top\mathbf{x} = \|\mathbf{x}\| > 0$ for all $\mathbf{x} \neq \mathbf{0}$, I (and thus $I+S$) is positive definite. Hence $I+S$ is always nonsingular by Problem **3**(a).

2. Let $A, B \in \mathbb{R}^{n\times n}$. Let $\lambda_a \in \mathbb{R}$ be an eigenvalue of A and $\lambda_b \in \mathbb{R}$ be an eigenvalue of B.

(a) Is it always true that $\lambda_a\lambda_b$ is an eigenvalue of AB? Is it always true that $\lambda_a + \lambda_b$ is an eigenvalue of $A+B$?
SOLUTION. No. $\lambda_a = \lambda_b = -1$ is an eigenvalue of both $A = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$ and $B = \left[\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$. However, $1 = \lambda_a\lambda_b$ is not an eigenvalue of $-I = AB$; also, $-2 = \lambda_a + \lambda_b$ is not an eigenvalue of $O = A+B$.

(b) Show that $\lambda \in \mathbb{R}$ is an eigenvalue of AB iff $\lambda \in \mathbb{R}$ is an eigenvalue of BA. [Hint: Homework **8**, Problem **1**(a).]
SOLUTION. Suppose $\lambda \neq 0$. Then by Homework **8**, Problem **1**(a), $\lambda I - AB = \lambda[I - (\lambda^{-1}A)B]$ is injective iff $\lambda I - BA = \lambda[I - B(\lambda^{-1}A)]$ is injective. So by Theorem **4.12**, we get $\ker(\lambda I - AB) \neq \{\mathbf{0}\}$ iff $\ker(\lambda I - BA) \neq \{\mathbf{0}\}$. That is, there exists $\mathbf{x} \neq \mathbf{0}$ such that $(\lambda I - AB)\mathbf{x} = \mathbf{0}$ iff there exists $\mathbf{y} \neq \mathbf{0}$ such that $(\lambda I - BA)\mathbf{y} = \mathbf{0}$. That is, there exists $\mathbf{x} \neq \mathbf{0}$ such that $AB\mathbf{x} = \lambda\mathbf{x}$ iff there exists $\mathbf{y} \neq \mathbf{0}$ such that $BA\mathbf{y} = \lambda\mathbf{y}$. That is, λ is an eigenvalue of AB iff λ is an eigenvalue of BA.
Suppose $\lambda = 0$. Let $\mathbf{x} \neq \mathbf{0}$ be a 0-eigenvector of AB, ie. $AB\mathbf{x} = 0\mathbf{x} = \mathbf{0}$. If $B\mathbf{x} \neq \mathbf{0}$, then we get $BA(B\mathbf{x}) = B(AB\mathbf{x}) = B\mathbf{0} = \mathbf{0} = 0(B\mathbf{x})$ and so $B\mathbf{x}$ is a 0-eigenvector of BA. If $B\mathbf{x} = \mathbf{0}$ and A is nonsingular, then $BA(A^{-1}\mathbf{x}) = B\mathbf{x} = \mathbf{0} = 0(A^{-1}\mathbf{x})$; note that $A^{-1}\mathbf{x} \neq \mathbf{0}$ and so $A^{-1}\mathbf{x}$ is a 0-eigenvector of BA. If A is singular, let $\mathbf{y}$ be a nonzero vector in $\operatorname{nullsp}(A)$, then $BA\mathbf{y} = B\mathbf{0} = \mathbf{0} = 0\mathbf{y}$ and so $\mathbf{y}$ is a 0-eigenvector of BA. In all cases, 0 is an eigenvalue of BA.

(c) Let $\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_d \in \mathbb{R}$. Show that

$$\alpha_0 + \alpha_1\lambda_a + \alpha_2\lambda_a^2 + \cdots + \alpha_d\lambda_a^d \in \mathbb{R}$$

is an eigenvalue of the matrix

$$\alpha_0 I + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_d A^d \in \mathbb{R}^{n\times n}.$$

SOLUTION. Let $\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n$ be a λ_a-eigenvalue of A. Then $A\mathbf{x} = \lambda_a\mathbf{x}$. So

$$\begin{aligned} A^2\mathbf{x} &= A(A\mathbf{x}) = A(\lambda_a\mathbf{x}) = \lambda_a A\mathbf{x} = \lambda_a^2\mathbf{x}, \\ A^3\mathbf{x} &= A(A^2\mathbf{x}) = A(\lambda_a^2\mathbf{x}) = \lambda_a^2 A\mathbf{x} = \lambda_a^3\mathbf{x}, \\ &\vdots \\ A^d\mathbf{x} &= A(A^{d-1}\mathbf{x}) = A(\lambda_a^{d-1}\mathbf{x}) = \lambda_a^{d-1}A\mathbf{x} = \lambda_a^d\mathbf{x}. \end{aligned}$$

Hence

$$
\begin{aligned}
(\alpha_0 I + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_d A^d)\mathbf{x} &= \alpha_0 \mathbf{x} + \alpha_1 A\mathbf{x} + \alpha_2 A^2\mathbf{x} + \cdots + \alpha_d A^d \mathbf{x} \\
&= \alpha_0 \mathbf{x} + \alpha_1 \lambda_a \mathbf{x} + \alpha_2 \lambda_a^2 \mathbf{x} + \cdots + \alpha_d \lambda_a^d \mathbf{x} \\
&= (\alpha_0 + \alpha_1 \lambda_a + \alpha_2 \lambda_a^2 + \cdots + \alpha_d \lambda_a^d)\mathbf{x}.
\end{aligned}
$$

(d) Show that if A is nonsingular, then $\lambda_a \neq 0$ and $1/\lambda_a$ is an eigenvalue of A^{-1}.
Solution. If 0 is an eigenvalue of A, then there exists $\mathbf{x} \neq \mathbf{0}$ such that $A\mathbf{x} = 0\mathbf{x} = \mathbf{0}$ and so $\operatorname{nullsp}(A) \neq \{\mathbf{0}\}$ and so A is not nonsingular by Theorem **4.12**. Hence $\lambda_a \neq 0$. Multiplying both sides of $A\mathbf{x} = \lambda_a \mathbf{x}$ by A^{-1} and dividing by λ_a then yields

$$\lambda_a^{-1}\mathbf{x} = A^{-1}\mathbf{x}.$$

3. A matrix $M \in \mathbb{R}^{n\times n}$ is called *positive semidefinite* if

$$\mathbf{x}^\top M \mathbf{x} \geq 0$$

for all $\mathbf{x} \in \mathbb{R}^n$. M is called *positive definite* if (i) M is positive semidefinite; and (ii) $\mathbf{x}^\top M\mathbf{x} = 0$ only if $\mathbf{x} = \mathbf{0}$.

(a) Show that every positive definite matrix is nonsingular (ie. invertible).
Solution. Let $\mathbf{x} \in \operatorname{nullsp}(M)$. So $M\mathbf{x} = \mathbf{0}$ and so $\mathbf{x}^\top M\mathbf{x} = \mathbf{x}^\top \mathbf{0} = 0$ and so $\mathbf{x} = \mathbf{0}$ since M is positive definite. Hence $\operatorname{nullsp}(M) = \{\mathbf{0}\}$ and M is nonsingular by Theorem **4.12**.

(b) Show that if M is positive semidefinite and $\lambda \in \mathbb{R}$ is an eigenvalue of M, then $\lambda \geq 0$.
Solution. Let $\mathbf{x} \in \mathbb{R}^n$ be the eigenvector of M corresponding to λ, ie. $M\mathbf{x} = \lambda\mathbf{x}$. So

$$\mathbf{x}^\top M\mathbf{x} = \lambda \mathbf{x}^\top \mathbf{x} = \lambda \|\mathbf{x}\|_2^2.$$

Since $\mathbf{x} \neq \mathbf{0}$ (eigenvectors are non-zero), so $\|\mathbf{x}\|_2 > 0$, and we get

$$\lambda = \frac{\mathbf{x}^\top M\mathbf{x}}{\|\mathbf{x}\|_2^2} \geq 0$$

since M is positive semidefinite.

(c) Show that if M is positive definite and $\lambda \in \mathbb{R}$ is an eigenvalue of M, then $\lambda > 0$.
Solution. Identical argument as above except that we have

$$\lambda = \frac{\mathbf{x}^\top M\mathbf{x}}{\|\mathbf{x}\|_2^2} > 0$$

since M is positive definite.

(d) Let M be positive definite and let

$$S_+ := \frac{1}{2}(M + M^\top) \quad \text{and} \quad S_- := \frac{1}{2}(M - M^\top).$$

Show that S_+ is a symmetric positive definite matrix and that

$$\mathbf{x}^\top M\mathbf{x} = \mathbf{x}^\top S_+ \mathbf{x}$$

for all $\mathbf{x} \in \mathbb{R}^n$. Show that S_- is a skew-symmetric matrix, and that

$$\mathbf{x}^\top S_- \mathbf{x} = 0$$

for all $\mathbf{x} \in \mathbb{R}^n$.
Solution. It is easy to check that

$$S_+^\top = \frac{1}{2}(M + M^\top)^\top = \frac{1}{2}(M^\top + M) = S_+$$

and that

$$S_-^\top = \frac{1}{2}(M - M^\top)^\top = \frac{1}{2}(M^\top - M) = -S_-,$$

and so S_+ is symmetric and S_- is skew-symmetric. Since any scalar (ie. 1×1 matrix) is equal to its own transpose,

$$\mathbf{x}^\top S_- \mathbf{x} = (\mathbf{x}^\top S_- \mathbf{x})^\top = \mathbf{x}^\top S_-^\top \mathbf{x} = -\mathbf{x}^\top S_- \mathbf{x},$$

implying that $\mathbf{x}^\top S_- \mathbf{x} = 0$ for any $\mathbf{x} \in \mathbb{R}^n$. Observe that,

$$M = \frac{1}{2}(M + M^\top) + \frac{1}{2}(M - M^\top) = S_+ + S_-.$$

Since $\mathbf{x}^\top S_- \mathbf{x} = 0$, we have

$$\mathbf{x}^\top M \mathbf{x} = \mathbf{x}^\top (S_+ + S_-)\mathbf{x} = \mathbf{x}^\top S_+ \mathbf{x} + \mathbf{x}^\top S_- \mathbf{x} = \mathbf{x}^\top S_+ \mathbf{x}$$

for all $\mathbf{x} \in \mathbb{R}^n$. Since M is positive definite,

$$\mathbf{x}^\top S_+ \mathbf{x} = \mathbf{x}^\top M \mathbf{x} > 0$$

for all non-zero $\mathbf{x} \in \mathbb{R}^n$. So S_+ is symmetric positive definite.

(e) Show that if M is symmetric positive definite, then $\langle \cdot , \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$,

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top M \mathbf{y},$$

is an inner product on $\mathbb{R}^n$.

Solution. *Positive definitness*: $\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^\top M \mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$ and $\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^\top M \mathbf{x} = 0$ implies that $\mathbf{x} = \mathbf{0}$. *Symmetry*: since M is symmetric, $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top M \mathbf{y} = \mathbf{x}^\top M^\top \mathbf{y} = (\mathbf{x}^\top M^\top \mathbf{y})^\top = \mathbf{y}^\top M \mathbf{x} = \langle \mathbf{y}, \mathbf{x} \rangle$. *Bilinearity*: $\langle \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2, \mathbf{y} \rangle = (\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2)^\top M \mathbf{y} = \alpha_1 \mathbf{x}_1^\top M \mathbf{y} + \alpha_2 \mathbf{x}_2^\top M \mathbf{y} = \langle \alpha_1 \mathbf{x}_1, \mathbf{y} \rangle + \langle \alpha_2 \mathbf{x}_2, \mathbf{y} \rangle$.

4. Let $A \in \mathbb{R}^{m \times n}$.

(a) Show that $A^\top A$ and $AA^\top$ are symmetric positive semidefinite matrices. Hence deduce that singular values are always nonnegative.

Solution. This just follows from the observation that

$$\mathbf{x}^\top A^\top A \mathbf{x} = (A\mathbf{x})^\top (A\mathbf{x}) = \|A\mathbf{x}\|_2^2 \geq 0$$

for all $\mathbf{x} \in \mathbb{R}^n$ and that

$$\mathbf{y}^\top A A^\top \mathbf{y} = (A^\top \mathbf{y})^\top (A^\top \mathbf{y}) = \|A^\top \mathbf{y}\|_2^2 \geq 0$$

for all $\mathbf{y} \in \mathbb{R}^m$.

(b) Show that if A is *full rank*, ie. $\operatorname{rank}(A) = \min\{m, n\}$, then either $A^\top A$ or $AA^\top$ must be positive definite.

Solution. If A is full-rank, then $\operatorname{rank}(A) = \min\{m, n\}$. If $m \geq n$, then $\operatorname{rank}(A) = n$. By the rank-nullity theorem,

$$\operatorname{nullity}(A) = n - \operatorname{rank}(A) = 0.$$

If $\mathbf{x}^\top A^\top A \mathbf{x} = 0$, then $\|A\mathbf{x}\|_2^2 = 0$; so $A\mathbf{x} = \mathbf{0}$; so $\mathbf{x} \in \operatorname{nullity}(A)$; so $\mathbf{x} = \mathbf{0}$. Hence $A^\top A$ is positive definite. On the other hand, if $m < n$, then $\operatorname{rank}(A) = m$; so $\operatorname{rank}(A^\top) = m$. By the rank-nullity theorem,

$$\operatorname{nullity}(A^\top) = m - \operatorname{rank}(A^\top) = 0.$$

If $\mathbf{y}^\top A A^\top \mathbf{y} = 0$, then $\|A^\top \mathbf{y}\|_2^2 = 0$; so $A^\top \mathbf{y} = \mathbf{0}$; so $\mathbf{y} \in \operatorname{nullity}(A^\top)$; so $\mathbf{y} = \mathbf{0}$. Hence $AA^\top$ is positive definite.

(c) Let $\lambda \in \mathbb{R}$ be an eigenvalue and $\mathbf{x} \in \mathbb{R}^{m+n}$ be a corresponding eigenvector of the matrix

$$\begin{bmatrix} O & A \\ A^\top & O \end{bmatrix} \in \mathbb{R}^{(m+n)\times(m+n)},$$

(written in block matrix form where O denotes a zero matrix of the appropriate size). Show that $\sigma = |\lambda|$ is a singular value of A and if $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^n$ are such that

$$\mathbf{x} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix},$$

then $\mathbf{u}$ is a left singular vector and $\mathbf{v}$ is a right singular vector of A corresponding to the singular value σ.

Solution. Since $\mathbf{x}$ is a λ-eigenvalue of the matrix,

$$\begin{bmatrix} O & A \\ A^\top & O \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \lambda \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}.$$

Using block matrix multiplication, we obtain

$$A\mathbf{v} = \lambda \mathbf{u}, \quad A^\top \mathbf{u} = \lambda \mathbf{v}.$$

Suppose $\lambda \neq 0$. Substituting the first equation into the second gives $A^\top(\lambda^{-1} A\mathbf{v}) = \lambda \mathbf{v}$ and so

$$A^\top A\mathbf{v} = \lambda^2 \mathbf{v}.$$

In other words, $\mathbf{v}$ is a right singular vector of A corresponding to the singular value $|\lambda| = \sqrt{\lambda^2}$. Substituting the second equation into the first gives $A(\lambda^{-1} A^\top \mathbf{u}) = \lambda \mathbf{u}$ and so

$$AA^\top \mathbf{u} = \lambda^2 \mathbf{u}.$$

In other words, $\mathbf{u}$ is a left singular vector of A corresponding to the singular value $|\lambda| = \sqrt{\lambda^2}$.

(d) Let $m = n$, ie. A is a square matrix. Show that

$$\frac{1}{2}\mathbf{x}^\top (A + A^\top)\mathbf{x} \leq (\mathbf{x}^\top A^\top A\mathbf{x})^{\frac{1}{2}}$$

for all $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_2 = 1$.

Solution. Note that $\mathbf{x}^\top A\mathbf{x} = (\mathbf{x}^\top A\mathbf{x})^\top = \mathbf{x}^\top A^\top \mathbf{x}$ and so

$$\begin{aligned} \frac{1}{2}\mathbf{x}^\top (A + A^\top)\mathbf{x} &= \frac{1}{2}(\mathbf{x}^\top A\mathbf{x} + \mathbf{x}^\top A^\top \mathbf{x}) \\ &= \mathbf{x}^\top A\mathbf{x} \\ &= \langle \mathbf{x}, A\mathbf{x} \rangle \\ &\leq \|\mathbf{x}\|_2 \|A\mathbf{x}\|_2, \end{aligned}$$

and since $\|\mathbf{x}\|_2 = 1$, we get

$$\frac{1}{2}\mathbf{x}^\top (A + A^\top)\mathbf{x} \leq \|A\mathbf{x}\|_2 = (\mathbf{x}^\top A^\top A\mathbf{x})^{\frac{1}{2}}.$$