The University of Texas at Austin
Department of Electrical and Computer Engineering

EE380K: Linear Systems Theory—Fall 2008

SOLUTIONS FOR PROBLEM SET TWO

Constantine Caramanis[1] Due: Wednesday, September 17, 2008.

---

1. Various properties of orthogonal subspaces: Let V be a finite dimensional vector space with an inner product, and let $U \subseteq V$ be a subspace. Recall that the space $U^\perp$ is defined as:

$$U^\perp = \{v \in V : \langle v, u\rangle = 0, \forall u \in U\}$$

(a) Show that if U is a subspace, the so is $U^\perp$:
Let $v_1, v_2 \in U^\perp$. $\forall u \in U, \langle v_1, u\rangle = \langle v_2, u\rangle = 0$. Hence,

$$\begin{aligned}\langle \alpha v_1 + \beta v_2, u\rangle &= \langle \alpha v_1, u\rangle + \langle \beta v_2, u\rangle \\ &= \alpha\langle v_1, u\rangle + \beta\langle v_2, u\rangle = 0 \\ &\Rightarrow (\alpha v_1 + \beta v_2) \in U^\perp.\end{aligned}$$

Also note that $\langle \mathbf{0}, u\rangle = 0 \Rightarrow \mathbf{0} \in U^\perp$. Other vector space properties are automatically inherited from V.
Hence, $U^\perp$ is a subspace.

(b) Show that $(U^\perp)^\perp = U$:
In class, several times I have used the notion of an orthonormal basis, i.e., a basis made of norm (measured with the 2-norm) 1 vectors, that are mutually orthogonal. If you have not seen it before, it is worthwhile seeing the Gram-Schmidt orthogonalization process. Here it is:

Gram-Schmidt

i. Given: A basis $\{v_1, \ldots, v_k\}$ for a subspace of an inner-product space V, so $U \subseteq V$, the goal is to find an orthonormal basis $\{\tilde{v}_1, \ldots, \tilde{v}_k\}$ for U.
ii. Let $\tilde{v}_1 = v_1/||v_1||_2$.
iii. Let $w_2 = v_2 - \langle v_2, \tilde{v}_1\rangle\tilde{v}_1$. Note that $w_2 \perp \tilde{v}_1$, since we have:

$$\langle w, \tilde{v}_1\rangle = \langle v_2, \tilde{v}_1\rangle - \langle v_2, \tilde{v}_1\rangle\langle \tilde{v}_1, \tilde{v}_1\rangle = 0.$$

Then, take $\tilde{v}_2 = w_2/||w_2||_2$.
iv. When $\tilde{v}_1, \ldots, \tilde{v}_i$ have been constructed, construct $\tilde{v}_{i+1}$ as follows. Let

$$w_{i+1} = v_{i+1} - \sum_{j=1}^{i}\langle v_{i+1}, \tilde{v}_j\rangle\tilde{v}_j,$$

and take $\tilde{v}_{i+1} = w_{i+1}/||w_{i+1}||_2$.

[1] Many solutions written in whole or in part by Johnson Carroll.

v. Check that this indeed gives an orthonormal basis.

Now let's use this to prove the claim: Take any basis for U. Using Gram-Schmidt if necessary, let us assume that this is an orthonormal basis: $\{v_1, \ldots, v_k\}$. Extend this (again, using G-S if required) to an orthonormal basis for all of V: $\{v_1, \ldots, v_k, u_1, \ldots, u_{n-k}\}$ (where V is taken to be n-dimensional). Now, it is easy to see that $\{u_1, \ldots, u_{n-k}\}$ is a basis for $U^\perp$. It is certainly independent. To show it spans, take any $v \in U^\perp$. We can always write $v = \sum_{i=1}^k \alpha_i v_i + \sum_{j=1}^{n-k} \beta_j u_j$. We have to show that $\alpha_i = 0$ for all i. To this end, we have: $0 = \langle v, v_i \rangle = \alpha_i$. But now the original claim follows, since $(U^\perp)^\perp$ must be spanned by the complement of the orthonormal basis elements that span $U^\perp$, and hence it is equal to U.

(c) Show that if $U, W \subseteq V$ are subspaces of V, then

$$U \subseteq W \Leftrightarrow U^\perp \supseteq W^\perp.$$

$$\begin{aligned}
U \subseteq W &\Leftrightarrow \forall u \in U, u \in W \\
&\Leftrightarrow \forall \hat{w} \in W^\perp \text{ and } u \in U, \langle \hat{w}, u \rangle = 0 \\
&\Leftrightarrow \forall \hat{w} \in W^\perp, \hat{w} \in U^\perp \\
&\Leftrightarrow W^\perp \subseteq U^\perp
\end{aligned}$$

(d) Suppose now that $X \subseteq V$ is just a subset, i.e., not necessarily a subspace of V. Show that the definition $X^\perp$ still makes sense, and that $X^\perp$ is a subspace. Next show that $(X^\perp)^\perp \supseteq X$, and it is defined as the smallest subspace that contains the set X.
$X^\perp$ is defined as:

$$X^\perp = \{v \in V : \langle v, x \rangle = 0, \forall x \in X\}$$

This definition produces no inconsistencies, and $X^\perp$ is a subspace; just as before, let $v_1, v_2 \in X^\perp$. $\forall x \in X, \langle v_1, x \rangle = \langle v_2, x \rangle = 0$. Hence,

$$\begin{aligned}
\langle \alpha v_1 + \beta v_2, x \rangle &= \langle \alpha v_1, x \rangle + \langle \beta v_2, x \rangle \\
&= \alpha \langle v_1, x \rangle + \beta \langle v_2, x \rangle = 0 \\
&\Rightarrow (\alpha v_1 + \beta v_2) \in X^\perp.
\end{aligned}$$

Also note that $\langle \mathbf{0}, x \rangle = 0 \Rightarrow \mathbf{0} \in X^\perp$. Other vector space properties are automatically inherited from V.
Hence, $X^\perp$ is a subspace.
Next, assume that $W \supseteq X$ is a subspace. Similarly to part (b),

$$\begin{aligned}
X \subseteq W &\Leftrightarrow \forall x \in X, x \in W \\
&\Rightarrow \forall \hat{w} \in W^\perp \text{ and } x \in X, \langle \hat{w}, x \rangle = 0 \\
&\Rightarrow \forall \hat{w} \in W^\perp, \hat{w} \in X^\perp \\
&\Rightarrow W^\perp \subseteq X^\perp \\
&\Leftrightarrow (X^\perp)^\perp \subseteq (W^\perp)^\perp = W \text{ by part (b)}
\end{aligned}$$

Since any subspace containing X also contains $(X^\perp)^\perp$, $(X^\perp)^\perp$ is the smallest subspace containing X.

(e) Show that V is the direct product of U and $U^\perp$ (denoted $V = U \oplus U^\perp$). That is, show that any $v \in V$ can be written *uniquely* as $v = u + u^\perp$, where $u \in U$, and $u^\perp \in U^\perp$.
Suppose $dim(V) = n$, $dim(U) = m \le n$. Let $\{u_i\}_{i=1}^m$ be an orthogonal basis of U, and extend the basis orthogonally $\{u_i\}_{i=m+1}^n$. Note that for any $\hat{u} \in U^\perp$, $\hat{u}$ can be a generically represented as linear combination of $\{u_i\}_{i=m+1}^n$. Also, for $\hat{u} = \sum a_i v_{i\,i=m+1}^{\,n}$, for any $u \in U$, $\langle \hat{u}, u \rangle = \langle \sum a_i v_{i\,i=m+1}^{\,n}, \sum b_i v_{i\,i=1}^{\,m} \rangle = 0$, since the basis is orthogonal and the inner product is bi-linear. Hence, $\{u_i\}_{i=m+1}^n$ is a basis for $U^\perp$.
For any $v \in V$, v can be written as a unique combination of basis elements:

$$
\begin{aligned}
v &= \sum a_i v_{i\,i=1}^{\,n} \\
&= \sum a_i v_{i\,i=1}^{\,n} + \sum a_i v_{i\,i=m+1}^{\,n}
\end{aligned}
$$

Define $u_v = \sum a_i v_{i\,i=1}^{\,n} \in U$ and $u_v^\perp = \sum a_i v_{i\,i=m+1}^{\,n} \in U^\perp$, and the result is shown.

2. Unitary Matrices. Recall from class that a square matrix U is unitary if its columns form an orthonormal set. This is equivalent to the condition that its rows form an orthonormal set. (Square U has orthonormal columns *iff* $U^TU = I = UU^T$ (easily proved) *iff* U^T has orthonormal columns.) In class we said that one property of a unitary matrix is that it preserves inner products, and in particular preserves the 2-norm of a vector: $\|Uv\|_2 = \|v\|_2$.

   Now, recall that matrices A and B are called *similar* if there is a similarity transformation taking A to B, i.e., there exists a matrix S with $A = S^{-1}BS$.

   If S is unitary, then we say that A and B are *unitarily equivalent*. Show that unitary equivalence preserves the Fromenius norm, i.e., $\|A\|_F = \|B\|_F$. (Hint: Recall that $\|A\|_F = \sum |a_{ij}|^2{}_{ij} = trace(A'A)$).

   Let $A = S^TBS$, where $S^T = S^{-1}$. $\|A\|_F^2 = trace(A^TA) = trace((S^TBS)^T(S^TBS)) = trace(S^TB^TSS^TBS) = trace(S^TB^TBS) = trace(SS^TB^TB) = trace(B^TB) = \|B\|_F^2$ so $\|A\|_F = \|B\|_F$.

3. In class we stated that for a matrix $A \in \mathbb{C}^{m \times n}$,

$$
\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|
$$

$$
\|A\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|
$$

   We showed the latter. Prove the former.

$$
\begin{aligned}
\|A\|_1 &= \sup_{\|x\|_1 = 1} \|Ax\|_1 \\
&= \sup_{\|x\|_1 = 1} \sum_j |(Ax)_j| \\
&= \sup_{sum\{|x_i|\} = 1} \sum_j |(Ax)_j|
\end{aligned}
$$

Consider x with $\|x\|_1 = 1$ such that $x_k = \alpha > 0$, $x_l = \beta > 0$, $\alpha + \beta \leq 1$.

$$\sum_j |(Ax)_j| = \alpha\|A_k\|_1 + \beta\|A_l\|_1 + \sum_{j\neq k,l} |(Ax)_j|,$$

where A_k is the k^{th} column of A. If $\|A_k\|_1 \geq \|A_l\|_1$, then we achieve a higher sum by making $x_k = \alpha + \beta$ and $x_l = 0$. Performing this comparison recursively through all of the columns, and we conclude that the maximizing x^* is:

$$x_i^* = \begin{cases} 1 & i = a \\ 0 & else \end{cases}$$

where a is the index of the column with the largest 1-norm. In the context of the matrix norm, this insight shows that

$$\begin{aligned} \|A\|_1 &= \max_{1\leq j\leq n} |Ax^*| \\ &= \max_{1\leq j\leq n} \|A_j\|_1 \\ &= \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}| \end{aligned}$$

4. Consider two matrices, $A, B \in \mathbb{C}^{n\times n}$.

   (a) Suppose that they are simultaneously diagonalizable, i.e., there exists some S such that $A = S^{-1}\Lambda_A S$, and $B = S^{-1}\Lambda_B S$, where Λ_A and Λ_B are diagonal matrices. Show that A and B commute.
   Note that diagonal matrices clearly commute. $AB = S^{-1}\Lambda_A SS^{-1}\Lambda_B S = S^{-1}\Lambda_A\Lambda_B S = S^{-1}\Lambda_B\Lambda_A S = S^{-1}\Lambda_B SS^{-1}\Lambda_A S = BA$.

   (b) Next, suppose that A and B are not quite diagonalizable, but they are simultaneously similar to upper triangular matrices, i.e., there exists and S such that $A = S^{-1}\Delta_A S$ and $B = S^{-1}\Delta_B S$, where Δ_A and Δ_B are upper triangular matrices. Show that every eigenvalue of $(AB - BA)$ must be zero.
   Remember that eigenvalues are preserved under similarity transform, and the eigenvalues of a triangular matrix are the diagonal entries. To see this, it is enough to show that the determinant of a diagonal matrix is the product of the diagonal elements. This one can show by induction. The base case, $n = 1$, is clear. Then the inductive hypothesis follows quickly by using expansion by minors to compute the determinant (go down the first column, not across the first row!). It is similarly straightforward (and you should make sure that you feel comfortable doing so) to show that the product (and sum) of two upper triangular matrices is again upper triangular.

$$\begin{aligned} (AB - BA) &= (S^{-1}\Delta_A SS^{-1}\Delta_B S - S^{-1}\Delta_A SS^{-1}\Delta_B S) \\ &= S^{-1}(\Delta_A\Delta_B - \Delta_B\Delta_A)S \end{aligned}$$

   The eigenvalues of $(AB - BA)$ are the diagonal elements of $\Delta_A\Delta_B - \Delta_B\Delta_A$. But the diagonal elements of $\Delta_A\Delta_B$ are the diagonal elements of $\Delta_B\Delta_A$. Hence, the diagonal elements of $\Delta_A\Delta_B - \Delta_B\Delta_A$ (and therefore the eigenvalues of $(AB - BA)$) are all zeros.

5. In class we discussed the least-squares problem of finding a value x to minimize $\|Ax-y\|_2$. Let us now consider a twist on this problem. Suppose that the value of A and y is not precisely known. Instead, let us suppose that while we see A and y, the true values are

$$A_t = A + \Delta A, y_t = y + \Delta y.$$

Further, suppose we know that $\|\Delta A \Delta y\|_F \leq 1$, where F denotes the Frobenius norm. Fix a solution x. How much worse does x do in the worst case, and in the case where there is no noise, i.e., $\Delta A = 0$ and $\Delta y = 0$? That is, compute

$$\max_{\|\Delta A \Delta y\|_F \leq 1} \|(A+\Delta A)x - (y+\Delta y)\|_2.$$

For $\|\Delta A \Delta y\|_F \leq 1$,

$$\begin{aligned}
\|(A+\Delta A)x - (y+\Delta y)\|_2 &= \|(Ax-y) + (\Delta Ax - \Delta y)\|_2 \\
&\leq \|(Ax-y)\|_2 + \|(\Delta Ax - \Delta y)\|_2 \\
&= \|(Ax-y)\|_2 + \|[\Delta A|\Delta y] \begin{bmatrix} x \\ -1 \end{bmatrix} \|_2 \\
&= \|(Ax-y)\|_2 + \|[\Delta A|\Delta y]\|_2 \cdot \| \begin{bmatrix} x \\ -1 \end{bmatrix} \|_2 \\
&\leq \|(Ax-y)\|_2 + \|[\Delta A|\Delta y]\|_F \cdot \| \begin{bmatrix} x \\ -1 \end{bmatrix} \|_2 \\
&\leq \|(Ax-y)\|_2 + \| \begin{bmatrix} x \\ -1 \end{bmatrix} \|_2 \\
&= \|(Ax-y)\|_2 + \sqrt{\|x\|_2^2 + 1}
\end{aligned}$$

Now we will craftily choose a $[\Delta A \Delta y]$ to satisfy the above inequality with equality, thereby achieving the maximum.

$$[\Delta A \Delta y] = \frac{u}{\sqrt{\|x\|_2^2+1}} \begin{bmatrix} x^T & -1 \end{bmatrix}$$

where

$$u = \begin{cases} \frac{Ax-y}{\|Ax-y\|} & Ax \neq y \\ \text{any unit norm vector} & Ax = y \end{cases}$$

Now $[\Delta A \Delta y]$ has rank 1, so $\|[\Delta A \Delta y]\|_F = \|[\Delta A \Delta y]\|_2 = 1$. Also,

$$\begin{aligned}
\|(Ax-y) + (\Delta Ax - \Delta y)\|_2 &= \|(Ax-y) + \frac{\frac{\|x\|^2}{\|Ax-y\|}}{\sqrt{\|x\|^2+1}} \cdot (Ax-y) + \frac{\frac{1}{\|Ax-y\|}}{\sqrt{\|x\|^2+1}} \cdot (Ax-y)\|_2 \\
&= (1 + \frac{\frac{\|x\|^2}{\|Ax-y\|}}{\sqrt{\|x\|^2+1}} + \frac{\frac{1}{\|Ax-y\|}}{\sqrt{\|x\|^2+1}}) \cdot \|(Ax-y)\|_2 \\
&= \|(Ax-y)\|_2 + \|\frac{\frac{\|x\|^2}{\|Ax-y\|}}{\sqrt{\|x\|^2+1}} \cdot (Ax-y) + \frac{\frac{1}{\|Ax-y\|}}{\sqrt{\|x\|^2+1}} \cdot (Ax-y)\|_2 \\
&= \|(Ax-y)\|_2 + \|(\Delta Ax - \Delta y)\|_2.
\end{aligned}$$

So, for the $[\Delta A \Delta y]$ above, the inequalities hold with equality. Since $\|(Ax-y)\|_2 + \sqrt{\|x\|_2^2+1}$ is independent of $[\Delta A \Delta y]$, we know the maximum is achieved.

6. **Exercise 3.1** The first and the third facts given in the problem are the keys to solve this problem, in addition to the fact that:

$$UA = \begin{pmatrix} R \\ 0 \end{pmatrix}.$$

Here note that R is a nonsingular, upper-triangular matrix so that it can be inverted. Now the problem reduces to show that

$$\hat{x} = arg \min_x \|y - Ax\|_2^2 = arg \min_x (y - Ax)'(y - Ax)$$

is indeed equal to

$$\hat{x} = R^{-1} y_1.$$

Let's transform the problem into the familiar form. We introduce an error e such that

$$y = Ax + e,$$

and we would like to minimize $\|e\|_2$ which is equivalent to minimizing $\|y - Ax\|_2$. Using the property of an orthogonal matrix, we have that

$$\|e\|_2 = \|Ue\|_2.$$

Thus with $e = y - Ax$, we have

$$\begin{aligned} \|e\|_2^2 &= \|Ue\|_2^2 = e'U'Ue = (U(y-Ax))'(U(y-Ax)) = \|Uy - UAx\|_2^2 \\ &= \left\| \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} R \\ 0 \end{pmatrix} x \right\|_2^2 = (y_1 - Rx)'(y_1 - Rx) + y_2' y_2. \end{aligned}$$

Since $\|y_2\|_2^2 = y_2' y_2$ is just a constant, it does not play any role in this minimization. Thus we would lik to have

$$y_1 - R\hat{x} = 0$$

and because R is an invertible matrix, $\hat{x} = R^{-1} y_1$.

7. **Exercise 4.6**. a) Suppose $A \in C_n^m$ has full column rank. Then QR factorization for A can be easily constructed from SVD:

$$A = U \begin{pmatrix} \Sigma_n \\ 0 \end{pmatrix} V'$$

where Σ_n is a $n \times n$ diagonal matrix with singular values on the diagonal. Lt $Q = U$ and $R = \Sigma_n V'$ and we get the QR factorization. Since Q is an orthogonal matrix, we can represent any $Y \in C_p^m$ as

$$Y = Q \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$$

Next

$$\|Y - AX\|_F^2 = \|Q \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} - Q \begin{pmatrix} R \\ 0 \end{pmatrix} X\|_F^2 = \|Q \begin{pmatrix} Y_1 - RX \\ Y_2 \end{pmatrix} \|_F^2$$

Denote

$$D = \begin{pmatrix} Y_1 - RX \\ Y_2 \end{pmatrix}$$

and note that multiplication by an orthogonal matrix does not change Frobenius norm of the matrix:

$$\|QD\|_F^2 = tr(D'Q'QD) = tr(D'D) = \|D\|_F^2$$

Since Frobenius norm squared is equal to sum of squares of all elements, square of the Frobenius norm of a block matrix is equal to sum of the squares of Frobenius norms of the blocks:

$$\| \begin{pmatrix} Y_1 - RX \\ Y_2 \end{pmatrix} \|_F^2 = \|Y_1 - RX\|_F^2 + \|Y_2\|_F^2$$

Since Y_2 block can not be affected by choice of X matrix, the problem reduces to minimization of $\|Y_1 - RX\|_F^2$. Recalling that R is invertible (because A has full column rank) the solution is

$$X = R^{-1}Y$$

b) Evaluate the expression with the pseudoinverse using the representations of A and Y from part a):

$$(A'A)^{-1} A'Y = \left(\begin{bmatrix} R' & 0 \end{bmatrix} Q'Q \begin{bmatrix} R \\ 0 \end{bmatrix} \right)^{-1} \begin{bmatrix} R' & 0 \end{bmatrix} Q'Q \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = R^{-1} (R')^{-1} \begin{bmatrix} R' & 0 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = R^{-1}Y_1$$

From 4.5 b) we know that if a matrix has a full column rank, $A^+ = (A'A)^{-1} A'$, therefore both expressions give the same solutions.

c)

$$\|Y - AX\|_F^2 + \|Z - BX\|_F^2 = \| \begin{pmatrix} Y \\ Z \end{pmatrix} - \begin{pmatrix} A \\ B \end{pmatrix} X\|_F^2$$

Since A has full column rank, $\begin{pmatrix} A \\ B \end{pmatrix}$ also has full column rank, therefore we can apply results from parts a) and b) to conclude that

$$X = \left(\begin{pmatrix} A \\ B \end{pmatrix}' \begin{pmatrix} A \\ B \end{pmatrix} \right)^{-1} \begin{pmatrix} A \\ B \end{pmatrix}' \begin{pmatrix} Y \\ Z \end{pmatrix} = (A'A + B'B)^{-1} (A'Y + B'Z)$$

8. **Exercise 4.7**.

(a) We are asked to show that $\mu_{\underline{\Delta}}(A) = \rho(A)$ when $\underline{\Delta} = \{\alpha I \,:\, \alpha \in \mathbb{C}\}$. Looking at the denominator in the definition of the structured singular value, we have:

$$\begin{aligned} \min_{\Delta \in \underline{\Delta}} \{\sigma_{\max}(\Delta) \,:\, \det(I - \Delta A) = 0\} &= \min_{\alpha \in \mathbb{C}} \{|\alpha| \,:\, \det(I - \alpha A) = 0\} \\ &= \min_{\alpha \in \mathbb{C}} \{|\alpha| \,:\, \det(\frac{1}{\alpha} I - A) = 0\} \\ &= \frac{1}{\lambda_{\max}(A)}, \end{aligned}$$

which is what we wanted to show.

(b) We are asked to show that when there are no constraints on the set $\underline{\Delta}$, then the structured singular value is exactly the maximum singular value.

First we show that $\sigma_{\max}(A) \geq \mu_{\underline{\Delta}}(A)$. Consider any Δ such that $\det(I - \Delta A) = 0$. This means that the null space of $(I - \Delta A)$ is nonemtpy, and hence there exists some nonzero vector x (we can take x to have Euclidean norm equal to 1) with $(I - \Delta A)x = 0$. But then we have:

$$\begin{aligned} 1 &= ||Ix||_2 \\ &= ||\Delta A x||_2 \\ &\leq ||\Delta||_2 ||A||_2 \\ &= \sigma_{\max}(\Delta)\sigma_{\max}(A), \end{aligned}$$

where the last inequality follows from the submultiplicative property of the induced norm, and the last equality follows from what we showed in class on Monday, namely, that $\sigma_{\max}(A) = ||A||_2$, for any matrix A. From this the inequality follows.

For the reverse inequality, we exhibit a specific Δ. To that end, we use the SVD of the matrix A. We have:

$$A = U \Sigma V'.$$

Then, let

$$\Delta = \frac{1}{\sigma_{\max}}(A) u_1 v_1'.$$

Note that this is a rank one matrix, since we have the *outer product* of u_1 and v_1, rather than the inner product which would give a scalar. You can check that $\sigma_{\max}(\Delta) = 1/\sigma_{\max}(A)$, and that therefore this shows $\sigma_{\max}(A) \leq \mu_{\underline{\Delta}}(A)$.

(c) For $\underline{\Delta} = \{\mathrm{diag}(\alpha_1, \ldots, \alpha_n) \,:\, \alpha_i \in \mathbb{C}\}$ we are asked to show:

$$\rho(A) \leq \mu_{\underline{\Delta}}(A) = \mu_{\underline{\Delta}}(D^{-1}AD) \leq \sigma_{\max}(D^{-1}AD),$$

where $D \in \{\mathrm{diag}(d_1, \ldots, d_n) \,:\, d_i > 0\}$.

The first inequality follows immediately from part (a), and from the fact that if we increase the feasible set in any minimization problem, the optimal value can only get smaller. Similarly, the last inequality follows from part (b). The equality in the middle follows from the fact that diagonal matrices commute, and hence $(I - \Delta D^{-1}AD) = (I - D^{-1}(\Delta A)D) = D^{-1}(I - \Delta A)D$. Finally recall that for any two (square) matrices A and B, $\det(AB) = \det(A) \cdot \det(B)$.