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

EE380K: Linear Systems Theory—Fall 2008

SOLUTIONS FOR PROBLEM SET THREE

Constantine Caramanis[1] Due: Wednesday, September 24, 2008.

1. Positive semidefinite matrices:

   (a) Characterize the set of positive semidefinite matrices in terms of their singular value decomposition. (Note that A is square, by definition).
   First, useful results about Hermitian matrices:

      i. Eigenvalues of Hermitian matrices are real.
      If λ is an eigenvalue of Hermitian matrix A, then for some x, $Ax = \lambda x$. Taking the Hermitian conjugate yields $x^H A = \lambda^* x^H$. Left-multiplying the first equation by x^H and right-multiplying the second equation by x yields

      $$\lambda x^H x = x^H A x = \lambda^* x^H x,$$

      so $\lambda = \lambda^*$, which means $\lambda \in \mathbb{R}$.

      ii. Eigenvectors corresponding to distinct eigenvalues of a Hermitian matrix are orthonormal.
      For any two distinct eigenvalues λ_1, λ_2, we can find normalized eigenvectors v_1, v_2 where $Av_1 = \lambda_1 v_1, Av_2 = \lambda_2 v_2$.

      $$\lambda_1 v_1^H v_2 = v_1^H A v_2 = \lambda_2 v_1^H v_2 \Rightarrow v_1^H v_2(\lambda_1 - \lambda_2) = 0 \Rightarrow v_1^H v_2 = 0$$

      Further, if $Av_1 = \lambda v_1, Av_2 = \lambda v_2, v_1 \neq v_2$ (repeated eigenvalue), then we can always, instead of v_2, choose an eigenvector $\hat{v}_2 = \frac{v_2 - \langle v_1 v_2 \rangle v_1}{\|v_2 - \langle v_1 v_2 \rangle v_1\|}$, the normalized part of v_2 orthogonal to v_1. Hence, we can always form an orthonormal set of eigenvectors for a Hermitian matrix.

   From the above results, Hermitian matrix A has an eigen-decomposition $Q\Lambda Q'$, where Q is the matrix of orthonormal eigenvectors (and therefore unitary) and Λ is the real diagonal matrix of eigenvalues. If A is positive semidefinite, each eigenvalue must be positive (if $\lambda_i < 0$, then $q_i^H A q_i = q_i^H Q \Lambda Q' q_i = \lambda_i \|q_i\|^2 = \lambda_i < 0$, violating positive semidefiniteness). So Λ is a diagonal matrix with non-negative entries.
   From the structure, the eigen-decomposition $Q\Lambda Q'$ *is a singular value decomposition* of A, with $U = V = Q$. Hence we can characterize the Hermitian positive semidefinite matrices as those matrices whose singular value decomposition can be written as an eigen-decomposition.

---

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

(b) Show that any Hermitian positive semidefinite matrix A has a square root, i.e., there exists some Y such that $A = Y'Y$. Explicitly construct such a Y in terms of the SVD of A.
Note that any diagonal matrix $\Lambda = diag(\lambda_1, \lambda_2, ..., \lambda_n)$ with non-negative entries has an intuitive square root given by a diagonal matrix $\Lambda^{1/2} = diag(\lambda_1^{1/2}, \lambda_2^{1/2}, ..., \lambda_n^{1/2})$.
Writing the SVD of A as described in part (a), and expanding:

$$A = Q\Lambda Q' = Q\Lambda^{1/2}\Lambda^{1/2}Q' = Q\Lambda^{1/2}Q' = (Q\Lambda^{1/2}Q')^2.$$

Choosing $Y = Q\Lambda^{1/2}$ yields the desired form.

2. Suppose that A and B have compatible dimension. Show that if

$$\|Bx\|_2 \le \|Ax\|_2, \ \forall x,$$

then there exists a matrix Y with $\|Y\|_2 \le 1$ and

$$B = YA.$$

$\|Bx\|_2 \le \|Ax\|_2$ for all x, and therefore we have to show that we can find a Y with $\|Y\|_2 \le 1$, such that $B = YA$.

Now, if A were invertible, we would obviously just take $Y = BA^{-1}$ and then $\|Y\|_2 \le 1$, would follow immediately from:

$$\|Yx\|_2 = \|B(A^{-1}x)\|_2 \le \|A(A^{-1}x)\|_2 = \|x\|,$$

so if $\|Yx\|_2 \le \|x\|_2$ for all x, then clearly $\|Y\|_2 \le 1$.

Since we don't have that, though, we have to do basically the same thing, by focusing on range spaces and $null^{\perp}$, which we can do via the SVD.

Let $B = U_B\Sigma_B V_B'$ be the SVD of B, and let $A = \hat{U}_A\hat{\Sigma}_A\hat{V}_A'$ be the reduced SVD of A[2].
Take $Y = U_B S_B V_B' \hat{V}_A \hat{\Sigma}_A^{-1} \hat{U}_A'$, and convince yourself that $YA = B$.

Note that $null(Y)^{\perp} \subseteq range(B)$.

Now, the claim is that *any* Y that satisfies $B = YA$, and $null(Y)^{\perp} \subseteq range(A)$, must have $\|Y\|_2 \le 1$.

Let y be the singular vector of norm 1, that corresponds to the largest singular value of Y. Then in particular, $\|Yy\|_2 = \|Y\|_2$.

Now, certainly, $y \in null(Y)^{\perp}$, by definition of the SVD for Y. But then, $y \in Range(A)$, and in particular, $y = Ag$ for some g.

---

[2] When $A = U_A\Sigma_A V_A'$ has rank r (and therefore r non-zero singular values), the reduced SVD of A is given by $A = \hat{U}_A\hat{\Sigma}_A\hat{V}_A'$ where $\hat{U}_A$ is given by the first r columns of U_A, $\hat{V}_A$ is given by the first r columns of V_A, and $\hat{\Sigma}_A$ is the $r \times r$ diagonal matrix of the singular values of A. That is, $\hat{\Sigma}_A = diag(\sigma_{A1}, \sigma_{A2}, ..., \sigma_{Ar})$. This allows us to write $\hat{\Sigma}_A^{-1}$.

And now we are in business, since we have:

$$\|Yy\|_2 = \|YAg\|_2 = \|Bg\|_2 \leq \|Ag\|_2 = \|y\|_2 = 1,$$

which is what we wanted to show.

3. Suppose that

$$\left\| \begin{pmatrix} X \\ A \end{pmatrix} \right\|_2 \leq \gamma.$$

   (a) Show that there exists a matrix Y with $\|Y\|_2 \leq 1$ such that

$$X = Y(\gamma^2 I - A'A)^{1/2}$$

Let's do some fancy norm manipulation. Let $v \in Domain(\begin{pmatrix} X \\ A \end{pmatrix})$.

$$\left\| \begin{pmatrix} X \\ A \end{pmatrix} v \right\|_2^2 \leq \left\| \begin{pmatrix} X \\ A \end{pmatrix} \right\|_2^2 \|v\|_2^2 \leq \gamma^2 \|v\|_2^2 = v' gamma^2 I v$$

and

$$\left\| \begin{pmatrix} X \\ A \end{pmatrix} v \right\|_2^2 = \left\| \begin{pmatrix} Xv \\ Av \end{pmatrix} \right\|_2^2 = v'X'Xv + v'A'Av.$$

Combining the two expressions,

$$\begin{aligned} v'X'Xv + v'A'Av &\leq v'\gamma^2 I v \\ v'X'Xv &\leq v'\gamma^2 I v - v'A'Av = v'(\gamma^2 I - A'A)v \\ \|Xv\|_2^2 &\leq \|(\gamma^2 I - A'A)^{1/2} v\|_2^2, \end{aligned}$$

so

$$\|Xv\|_2 \leq \|(\gamma^2 I - A'A)^{1/2} v\|_2$$

for any v.
Hence, we can apply the results of problem (3) and conclude that there exists a matrix Y satisfying the desired properties.

   (b) Now show that if

$$\|\begin{pmatrix} X & A \end{pmatrix}\|_2 \leq \gamma.$$

then there exists a matrix Z such that $\|Z\|_2 \leq 1$, and

$$X = (\gamma^2 I - AA')^{1/2} Z$$

Note that for a matrix $C = U\Sigma V'$, the SVD of C' is given by $C' = V\Sigma' U'$. Since the 2-norm of a matrix is the largest singular value and Σ' has the same non-zero entries as Σ, $\|A\|_2 = \|A'\|_2$.
Applying this to the current problem,

$$\|\begin{pmatrix} X & A \end{pmatrix}\|_2 = \left\| \begin{pmatrix} X' \\ A' \end{pmatrix} \right\|_2 \leq \gamma.$$

We can apply the results of part (a) to find a matrix Z' with $\|Z'\|_2 \leq 1$ such that

$$X' = Z'(\gamma^2 I - A'A)^{1/2}.$$

Taking the conjugate transpose of both sides yields

$$X = (\gamma^2 I - A'A)^{1/2} Z,$$

and $\|Z\|_2 = \|Z'\|_2 \leq 1$, proving the result.

4. In this problem, you will prove the following result:

$$\min_X \left\| \begin{pmatrix} X & B \\ C & A \end{pmatrix} \right\|_2 = \max \left\{ \left\| \begin{pmatrix} C & A \end{pmatrix} \right\|_2, \left\| \begin{pmatrix} B \\ A \end{pmatrix} \right\|_2 \right\}$$

(a) Let γ_0 denote the value of the minimization of the left hand side above. Let γ_1 denote the right hand side. Show that for any choice of X,

$$\left\| \begin{pmatrix} X & B \\ C & A \end{pmatrix} \right\|_2 \geq \max \left\{ \left\| \begin{pmatrix} C & A \end{pmatrix} \right\|_2, \left\| \begin{pmatrix} B \\ A \end{pmatrix} \right\|_2 \right\}$$

$$\left\| \begin{pmatrix} X & B \\ C & A \end{pmatrix} \right\|_2 = \sup_{\|v\|_2 = 1} \left\| \begin{pmatrix} X & B \\ C & A \end{pmatrix} v \right\|_2$$

Partitioning v into appropriately sized parts,

$$\sup_{\|v\|_2=1} \left\| \begin{pmatrix} X & B \\ C & A \end{pmatrix} v \right\|_2 = \sup_{\|v\|_2=1} \left\| \begin{pmatrix} X & B \\ C & A \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right\|_2 = \sup_{\|v\|_2=1} \left\| \begin{pmatrix} Xv_1 + Bv_2 \\ Cv_1 + Av_2 \end{pmatrix} \right\|_2$$

If we let v_2 be the largest singular vector of $\begin{pmatrix} B \\ A \end{pmatrix}$ and let $v_1 = 0$, then $\|v\| = 1$ and

$$\left\| \begin{pmatrix} X & B \\ C & A \end{pmatrix} v \right\|_2 = \left\| \begin{pmatrix} B \\ A \end{pmatrix} v_2 \right\|_2 = \left\| \begin{pmatrix} B \\ A \end{pmatrix} \right\|_2$$

so

$$\left\| \begin{pmatrix} X & B \\ C & A \end{pmatrix} \right\|_2 \geq \left\| \begin{pmatrix} B \\ A \end{pmatrix} \right\|_2 .$$

Then,

$$\left\| \begin{pmatrix} X & B \\ C & A \end{pmatrix} \right\|_2 = \left\| \begin{pmatrix} X' & C' \\ B' & A' \end{pmatrix} \right\|_2 \geq \left\| \begin{pmatrix} C' \\ A' \end{pmatrix} \right\|_2 = \left\| \begin{pmatrix} C & A \end{pmatrix} \right\|_2$$

Combining the two inequalities into a maximum yields the result.

(b) Now, using the previous exercise, show that there are two matrices, Y and Z, with $\|Y\|_2 \leq 1$ and $\|Z\|_2 \leq 1$ such that

$$B = Y(\gamma_1^2 I - A'A)^{1/2}$$
$$C = (\gamma_1^2 I - AA')^{1/2} Z$$

Since

$$\left\| \begin{pmatrix} B \\ A \end{pmatrix} \right\|_2 \leq \gamma_1, \ \left\| \begin{pmatrix} C & A \end{pmatrix} \right\|_2 \leq \gamma_1,$$

the results of the previous problem can be applied directly to produce the required result.

(c) Let $X = -YA'Z$, and show that for this X, we have

$$\left\|\begin{pmatrix} X & C \\ B & A \end{pmatrix}\right\|_2 = \left\|\begin{pmatrix} Y & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} -A' & (\gamma_1^2 I - A'A)^{1/2} \\ (\gamma_1^2 I - AA')^{1/2} & A \end{pmatrix}\begin{pmatrix} Y & 0 \\ 0 & I \end{pmatrix}\right\|_2$$

Let's substitute for X, B, and C!

$$\begin{aligned}
\left\|\begin{pmatrix} X & B \\ C & A \end{pmatrix}\right\|_2 &= \left\|\begin{pmatrix} -YA'Z & Y(\gamma_1^2 I - A'A)^{1/2} \\ (\gamma_1^2 I - AA')^{1/2} Z & A \end{pmatrix}\right\|_2 \\
&= \left\|\begin{pmatrix} Y & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} -A'Z & (\gamma_1^2 I - A'A)^{1/2} \\ (\gamma_1^2 I - AA')^{1/2} Z & A \end{pmatrix}\right\|_2 \\
&= \left\|\begin{pmatrix} Y & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} -A' & (\gamma_1^2 I - A'A)^{1/2} \\ (\gamma_1^2 I - AA')^{1/2} & A \end{pmatrix}\begin{pmatrix} Z & 0 \\ 0 & I \end{pmatrix}\right\|_2,
\end{aligned}$$

the desired equality. Continuing with much algebra,

$$\begin{aligned}
\min_X \left\|\begin{pmatrix} X & B \\ C & A \end{pmatrix}\right\|_2^2 &\le \left\|\begin{pmatrix} -YA'Z & B \\ C & A \end{pmatrix}\right\|_2^2 \\
&= \left\|\begin{pmatrix} -YA'Z & Y(\gamma_1^2 I - A'A)^{1/2} \\ (\gamma_1^2 I - AA')^{1/2} Z & A \end{pmatrix}\right\|_2^2 \\
&= \left\|\begin{pmatrix} Y & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} -A'Z & (\gamma_1^2 I - A'A)^{1/2} \\ (\gamma_1^2 I - AA')^{1/2} Z & A \end{pmatrix}\right\|_2^2 \\
\le \left\|\begin{pmatrix} Y & 0 \\ 0 & I \end{pmatrix}\right\|_2^2 & \left\|\begin{pmatrix} -A' & (\gamma_1^2 I - A'A)^{1/2} \\ (\gamma_1^2 I - AA')^{1/2} & A \end{pmatrix}\right\|_2^2 \left\|\begin{pmatrix} Z & 0 \\ 0 & I \end{pmatrix}\right\|_2^2 \\
&\le \left\|\begin{pmatrix} -A' & (\gamma_1^2 I - A'A)^{1/2} \\ (\gamma_1^2 I - AA')^{1/2} & A \end{pmatrix}\right\|_2^2 \\
&= \sup_{\|v\| \le 1} \left\|\begin{pmatrix} -A' & (\gamma_1^2 I - A'A)^{1/2} \\ (\gamma_1^2 I - AA')^{1/2} & A \end{pmatrix} v\right\|_2^2
\end{aligned}$$

$$\begin{aligned}
&= \sup_{\|v\| \le 1} v' \begin{pmatrix} -A' & (\gamma_1^2 I - A'A)^{1/2} \\ (\gamma_1^2 I - AA')^{1/2} & A \end{pmatrix}' \begin{pmatrix} -A' & (\gamma_1^2 I - A'A)^{1/2} \\ (\gamma_1^2 I - AA')^{1/2} & A \end{pmatrix} v \\
&= \sup_{\|v\| \le 1} v' \begin{pmatrix} -A & (\gamma_1^2 I - AA')^{1/2} \\ (\gamma_1^2 I - A'A)^{1/2} & A' \end{pmatrix} \begin{pmatrix} -A' & (\gamma_1^2 I - A'A)^{1/2} \\ (\gamma_1^2 I - AA')^{1/2} & A \end{pmatrix} v \\
&= \sup_{\|v\| \le 1} v' \begin{pmatrix} AA' + \gamma_1^2 I - AA' & -A(\gamma_1^2 I - A'A) + (\gamma_1^2 I - AA')A \\ -(\gamma_1^2 I - A'A)A' + A'(\gamma_1^2 I - AA') & A'A + \gamma_1^2 I - A'A \end{pmatrix} v \\
&= \sup_{\|v\| \le 1} v' \begin{pmatrix} \gamma_1^2 I & 0 \\ 0 & \gamma_1^2 I \end{pmatrix} v \\
&= \sup_{\|v\| \le 1} v' \gamma_1^2 I v = \sup_{\|v\| \le 1} \gamma_1^2 v'v = \gamma_1^2.
\end{aligned}$$

So

$$\gamma_0 = \min_X \left\|\begin{pmatrix} X & B \\ C & A \end{pmatrix}\right\|_2 \le \gamma_1.$$

5. For $y = Ax + w$, $A \in \mathbb{R}^{m\times n}$, $m > n$, $rank(A) = n$, we can write the reduced SVD of A as described earlier: $A = \hat{U}_A\hat{\Sigma}_A\hat{V}_A'$, where $\hat{U}_A \in \mathbb{R}^{m\times n}$, $\hat{\Sigma}_A, \hat{V}_A \in \mathbb{R}^{n\times n}$. One possible B (an important one, as we shall see) is given by $B = \hat{V}_A\hat{\Sigma}_A^{-1}\hat{U}_A'$, and indeed $BA = \hat{V}_A\hat{\Sigma}_A^{-1}\hat{U}_A'\hat{U}_A\hat{\Sigma}_A\hat{V}_A' = I$.

   (a) Describe the set ϵ_α of all possible estimation errors, when the noise is norm-bounded, i.e., when $\|w\|_2 \leq \alpha$. What does it look like, with respect to the singular values and vectors of B?

   $$\{e = (x - By) : B \text{ is unbiased linear}\} = \{e = x - BAx - Bw : BA = I\} = \{e = -Bw\}$$

   $\{w : \|w\| \leq \alpha\}$ is the closed α-sphere in $\mathbb{R}^n$. Since $-B = U_B(-\Sigma_B)V_B'$, the set ϵ_α is a hyperellipse in $\mathbb{R}^m$ where the principle axes are given by the vectors $\{-\alpha\sigma_{Bi}u_{Bi}\}_{i=1}^n$, the images of the vectors $\{\alpha v_{Bi}\}_{i=1}^n$ under $-B$. ($\{u_{Bi}\}_{i=1}^n$ are the first n columns of U_B; similarly for $\{v_{Bi}\}_{i=1}^n$.)

   (b) What is the maximum $\| \ \|_2$-norm of the error?
   The maximal error is realized at the most extreme point of the hyperellipse, or via the first singular vector of $-B$: v_{B1}.
   $\| - B(\alpha v_{B1})\|_2 = \| - \alpha U_B\Sigma_B V_B v_{B1}\|_2 = \alpha\sigma_{B1}$.
   Using the properties of norms, $\|-Bw\|_2 \leq \|B\|_2\|w\|_2 \leq \|B\|_2\alpha = \sigma_{B1}\alpha$, which is realized with equality when w is parallel to the first column of V_B with norm α.

   (c) Let $B_{ls} = A^\dagger = (A^TA)^{-1}A^T = \hat{V}_A\hat{\Sigma}_A^{-1}\hat{U}_A'$ be the least-squares estimator. Show that B_{ls} is optimal in the sense that it minimizes the maximum error norm.
   Let B be any other estimator s.t. $BA = I$, and define $Z = B - B_{ls}$. Note that $ZA = Z\hat{U}_A\hat{\Sigma}_A\hat{V}_A' = 0$, so $Z\hat{U}_A = 0$. Therefore

   $$\sigma_{B1}^2 = \|B\|_2^2 = \|B'\|_2^2 = \sup_{\|x\|_2=1} x'BB'x.$$

   But for all x,

   $$\begin{aligned} x'BB'x &= x'(B_{ls} + Z)(B_{ls} + Z)'x \\ &= x'(B_{ls}B_{ls}' + B_{ls}Z' + B_{ls}'Z + ZZ')x \\ &= x'(B_{ls}B_{ls}' + ZZ')x \geq x'B_{ls}B_{ls}'x \end{aligned}$$

   using $ZB_{ls}' = (Z\hat{U})\hat{\Sigma}^{-1}\hat{V}' = 0$.
   So $\sigma_{B_{ls}1} = \|B_{ls}\|_2 \leq \|B\|_2 = \sigma_{B1}$, and we conclude that B_{ls} minimizes the maximum error norm.

6. Low Rank Approximation. For a matrix A with rank k, find the matrix $\hat{A}$ with rank $p < k$ that minimizes $\|A - \hat{A}\|_2$.

   Using SVD, A can be decomposed as

   $$A = U \begin{pmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_{p+1} & & & \\ & & & \ddots & & \\ & & & & \sigma_k & \\ & & & & & 0 \end{pmatrix} V',$$

where U and V are unitary matrices. Following the given procedure, let's select the first p+1 columns of V : $\{v_1, v_2, \cdots, v_{p+1}\}$. From the definition of SVD those v_i's are orthonormal, and hence independent. Note that $\{v_1, v_2, \cdots, v_{p+1}, \cdots v_n\}$ span $\mathbb{R}^n$, and if $rank(\hat{A}) = p$, then exactly p of the vectors, $\{v_1, v_2, \cdots, v_{p+1}, \cdots v_n\}$, span $\mathcal{R}(\hat{A}') = \mathcal{N}^{\perp}(\hat{A})$. The remaining vectors span $\mathcal{N}(\hat{A})$. So, given any $p+1$ linearly independent vectors in $\mathbb{R}^n$, at least one must be in the nullspace of $\hat{A}$. That is, $\exists \alpha_i$ for $i = 1, \cdots, p+1$ not all zero such that

$$\hat{A}(\alpha_1 v_1 + \alpha_2 v_2 + \cdots \alpha_{p+1} v_{p+1}) = 0.$$

This implies that there exists a nonzero vector z such that

$$z = \sum_{i=1}^{p+1} \alpha_i v_i = \begin{pmatrix} v_1 & \cdots & v_{p+1} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{p+1} \end{pmatrix}$$

with $\|z\|_2 = 1$ such that $\hat{A}z = 0$. Thus,

$$(A - \hat{A})z = Az = U\Sigma \begin{pmatrix} - & {v_1}' & - \\ & \vdots & \\ - & {v_{p+1}}' & - \end{pmatrix} \left(\sum_{i=1}^{p+1} \alpha_i v_i \right) = U \begin{pmatrix} \sigma_1 \alpha_1 \\ \sigma_2 \alpha_2 \\ \vdots \\ \sigma_{p+1} \alpha_{p+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{1}$$

By taking 2-norm of both sides of the above equation,

$$\begin{aligned} \|(A - \hat{A})z\|_2 &= \|U \begin{pmatrix} \sigma_1 \alpha_1 \\ \sigma_2 \alpha_2 \\ \vdots \\ \sigma_{p+1} \alpha_{p+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \|_2 = \| \begin{pmatrix} \sigma_1 \alpha_1 \\ \sigma_2 \alpha_2 \\ \vdots \\ \sigma_{p+1} \alpha_{p+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \|_2 \quad \text{(since U is a unitary matrix)} \\ &= \left(\sum_{i=1}^{p+1} |\sigma_i \alpha_i|^2 \right)^{1/2} \geq \sigma_{p+1} \left(\sum_{i=1}^{p+1} |\alpha_i|^2 \right)^{1/2}. \end{aligned} \tag{2}$$

But, from our construction of z,

$$\|z\|_2^2 = 1 \rightarrow \| \begin{pmatrix} v_1 & \cdots & v_{p+1} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{p+1} \end{pmatrix} \|_2^2 = 1 \rightarrow \| \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{p+1} \end{pmatrix} \|_2^2 = \sum_{i=1}^{p+1} |\alpha_i|^2 = 1.$$

Thus, equation(2) becomes

$$\|(A - \hat{A})z\|_2 \geq \sigma_{p+1}.$$

Finally, $\|(A-\hat{A})z\|_2 \leq \|A-\hat{A}\|_2$ for all z such that $\|z\|_2 = 1$.

Now we have to achieve the lower bound. Choose

$$\hat{A} = U \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{pmatrix} V'.$$

$\hat{A}$ has rank p, and

$$A - \hat{A} = U \begin{pmatrix} 0 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & & & \\ & & & \sigma_{p+1} & & & \\ & & & & \ddots & & \\ & & & & & \sigma_k & \\ & & & & & & 0 \end{pmatrix} V'.$$

where $\|A-\hat{A}\|_2 = \sigma_{p+1}$.

7. State whether each statement is true or false, and justify.

   (a) The statement is false, to construct a counterexample use $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

   (b) True, $\sigma_{max}(E) = \|E\|_2 = \|(E+A)-A\|_2$. $(E+A) = U\Sigma V'$. For $A, E \in \mathbb{R}^{m\times n}$, if A has rank $p < m, n$ then, as in the previous problem, there is a unit vector z constructed from the first $p+1$ columns of V such that $Az = 0$. Following the same procedure as before, $\|(E+A)-A\|_2 \geq \|((E+A)-A)z\|_2 = \|(E+A)z\|_2 \geq \sigma_{p+1}(E+A) \geq \sigma_{min}(E+A)$. Combining expressions yields the desired result.

   (c) True, first note that for any invertible matrix A we have $\sigma_{max}(A^{-1}) = \sigma_{min}(A)$, infact one can easily show that if $A \in \mathbf{C^{mxm}}$ then $\sigma_i(A) = \sigma_j(A^{-1})$ with $j = m - i + 1$ and $i \in \{1, ..., m\}$. Given that remark all we need to show is that $1 - \sigma_{max}(A) \leq \sigma_{min}(I - A)$. From the triangle inequality one has :

   $$|\|Ix\|_2 - \|Ax\|_2| \leq \|(I-A)x\|_2$$

   restricting our attention to unit vectors x, so $\|x\|_2 = 1$ and taking into account that $\sigma_{max}(A) < 1$ we obtain:

   $$1 - \|Ax\|_2 \leq \|(I-A)x\|_2$$

   but for any x on the unit sphere $1 - \sigma_{max}(A) \leq 1 - \|Ax\|_2$, thus

   $$1 - \sigma_{max}(A) \leq \|(I-A)x\|_2$$

the above equation is still valid for any x on the unit sphere, ie $\|x\|_2 = 1$ so we can choose x such that $\|(I-A)x\|_2 = \sigma_{min}(I-A)$, thus

$$1 - \sigma_{max}(A) \leq \sigma_{min}(I-A)$$

completing the proof.

(d) Consider a matrix E with $rank(E) \leq i-1$, but otherwise arbitrary. From the triangle inequality we have

$$\|A+B-E\|_2 \leq \|A-E\|_2 + \|B\|_2$$

from the previous exercise we know that $\sigma_i(A) \leq \|A-E\|_2$ if $rank(E) = i-1$, thus

$$\sigma_i(A+B) \leq \|A-E\|_2 + \sigma_{max}(B)$$

now choose the particular E that achieves the bound, ie. $\|A-E\|_2 = \sigma_i(A)$ and we get

$$\sigma_i(A+B) \leq \sigma_i(A) + \sigma_{max}(B)$$

by letting $B = I$ we obtain the desired result.

8. $*$ For the 2-norm, we can do the problem as follows: if A is invertible, it must have strictly positive singular values. In particular, its smallest singular value must be strictly positive, say, $\sigma_{\min} = \gamma > 0$. But then from our work on matrix perturbation, we know that the minimum 2-norm matrix Δ such that $(A+\Delta)$ is singular, must have norm γ. Therefore any matrix $\tilde{A}$ such that $||A-\tilde{A}||_2 < \gamma$, must itself be invertible.

   There is a more general approach that works for other norms as well: note that the determinant is a continuous map in the topology induced by the norm (in fact this topology is the same for all norms listed). Then recall that the inverse of a continuous function maps open sets to open sets. The set $\mathbb{R} \setminus \{0\}$ is open, and therefore the image under the inverse of the determinant map is again an open set. But this latter set is indeed the set of all invertible matrices, which is what we wanted to show.