Exercises

Exercise 4.1 Verify that for any A, an $m \times n$ matrix, the following holds:

$$\frac{1}{\sqrt{n}}\|A\|_1 \le \|A\|_2 \le \sqrt{m}\|A\|_\infty.$$

Exercise 4.5 Suppose the $m \times n$ matrix A is decomposed in the form

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

where U and V are unitary matrices, and Σ is an invertible $r \times r$ matrix (— the SVD could be used to produce such a decomposition). Then the "Moore-Penrose inverse", or *pseudo-inverse* of A, denoted by A^+, can be defined as the $n \times m$ matrix

$$A^+ = V \begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U'$$

(You can invoke it in Matlab with `pinv(A)`.)

a) Show that A^+A and AA^+ are symmetric, and that $AA^+A = A$ and $A^+AA^+ = A^+$. (These four conditions actually constitute an alternative definition of the pseudo-inverse.)

b) Show that when A has full column rank then $A^+ = (A'A)^{-1}A'$, and that when A has full row rank then $A^+ = A'(AA')^{-1}$.

c) Show that, of all x that minimize $\|y - Ax\|_2$ (and there will be many, if A does not have full column rank), the one with smallest length $\|x\|_2$ is given by $\hat{x} = A^+y$.

Exercise 4.7 Structured Singular Values

Given a complex square matrix A, define the *structured singular value function* as follows.

$$\mu_{\underline{\Delta}}(A) = \frac{1}{\min_{\Delta \in \underline{\Delta}}\{\sigma_{max}(\Delta) \mid \det(I - \Delta A) = 0\}}$$

where $\underline{\Delta}$ is some set of matrices.

a) If $\underline{\Delta} = \{\alpha I \, : \, \alpha \in \mathbb{C}\}$, show that $\mu_{\underline{\Delta}}(A) = \rho(A)$, where ρ is the *spectral radius* of A, defined as: $\rho(A) = \max_i |\lambda_i|$ and the λ_i's are the eigenvalues of A.

b) If $\underline{\Delta} = \{\Delta \in \mathbb{C}^{n \times n}\}$, show that $\mu_{\underline{\Delta}}(A) = \sigma_{max}(A)$

c) If $\underline{\Delta} = \{diag(\alpha_1, \cdots, \alpha_n) \mid \alpha_i \in \mathbb{C}\}$, show that

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

where

$$D \in \{\text{diag}(d_1, \cdots, d_n) \mid d_i > 0\}$$

Exercise 4.8 Consider again the *structured singular value function* of a complex square matrix A defined in the preceding problem. If A has more structure, it is sometimes possible to compute $\mu_{\underline{\Delta}}(A)$ exactly. In this problem, assume A is a rank-one matrix, so that we can write $A = uv'$ where u, v are complex vectors of dimension n. Compute $\mu_{\underline{\Delta}}(A)$ when

(a) $\underline{\Delta} = diag(\delta_1, \ldots, \delta_n), \quad \delta_i \in \mathbb{C}$.

(b) $\underline{\Delta} = diag(\delta_1, \ldots, \delta_n), \quad \delta_i \in \mathbb{R}$.

To simplify the computation, minimize the Frobenius norm of Δ in the defintion of $\mu_{\underline{\Delta}}(A)$.

Exercise 5.1 Suppose the complex $m \times n$ matrix A is perturbed to the matrix $A + E$.

(a) Show that

$$|\,\sigma_{max}(A+E) \;-\; \sigma_{max}(A)\,| \;\leq\; \sigma_{max}(E)$$

Also find an E that results in the inequality being achieved with equality.

(Hint: To show the inequality, write $(A + E) = A + E$ and $A = (A + E) - E$, take the 2-norm on both sides of each equation, and use the triangle inequality.)

It turns out that the result in (a) actually applies to *all* the singular values of A and $A+E$, not just the largest one. Part (b) below is one version of the result for the smallest singular value.

(b) Suppose A has *less than* full column rank, i.e. has rank$< n$, but $A + E$ has full column rank. Show (following a procedure similar to part (a) — but looking at $\min \|(A + E)x\|_2$ rather than the norm of $A + E$, etc.) that

$$\sigma_{min}(A+E) \;\leq\; \sigma_{max}(E)$$

Again find an E that results in the inequality being achieved with equality.

[The result in (b), and some extensions of it, give rise to the following sound (and widely used) procedure for estimating the rank of some underlying matrix A, given only the matrix $A + E$ and knowledge of $\|E\|_2$: Compute the SVD of $A + E$, then declare the "numerical rank" of A to be the number of singular values of $A + E$ that are larger than the threshold $\|E\|_2$. The given information is consistent with having an A of this rank.]

(c) Verify the above results using your own examples in MATLAB. You might also find it interesting to verify numerically that for large m, n, the norm of the matrix $E = s * randn(m, n)$ — which is a matrix whose entries are independent, zero-mean, Gaussian, with standard deviation s — is close to $s * (\sqrt{m} + \sqrt{n})$. So if A is perturbed by such a matrix, then a reasonable value to use as a threshold when determining the numerical rank of A is this number.