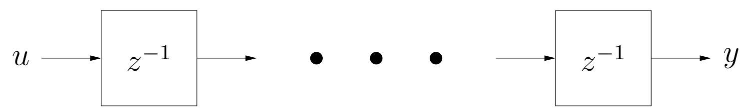
EE263 homework 8 solutions

13.17 *FIR filter with small feedback.* Consider a cascade of 100 one-sample delays:

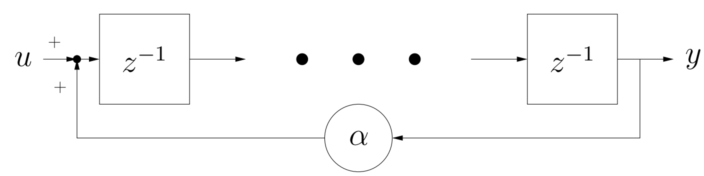


(a) Express this as a linear dynamical system

$$x(t+1) = Ax(t) + Bu(t), \qquad y(t) = Cx(t) + Du(t)$$

(b) What are the eigenvalues of A?

(c) Now we add simple feedback, with gain $\alpha = 10^{-5}$, to the system:



Express this as a linear dynamical system

$$x(t+1) = A_f x(t) + B_f u(t), \qquad y(t) = C_f x(t) + D_f u(t)$$

(d) What are the eigenvalues of A_f?

(e) How different is the impulse response of the system with feedback ($\alpha = 10^{-5}$) and without feedback ($\alpha = 0$)?

Solution. For the system without feedback, we can take

$$A = \begin{bmatrix} 0 & \cdots & & & \\ 1 & 0 & \cdots & & \\ 0 & 1 & 0 & \cdots & \\ & & \ddots & & \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}, \quad D = 0.$$

The eigenvalues of A are all zero. In fact, A has only one Jordan block.

The impulse response of this system is very simple. It's nothing more than a single unit impulse at $t = 100$. The system simply delays the input signal 100 samples.

Now we add feedback. For this system we can take

$$A = \begin{bmatrix} 0 & \cdots & & \cdots & 10^{-5} \\ 1 & 0 & \cdots & & \\ 0 & 1 & 0 & \cdots & \\ & & \ddots & & \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}, \quad D = 0.$$

The only difference is the entry 10^{-5} in the upper righthand corner.

Let's find the eigenvalues of A. We can use some determinant formulas, or just look at the poles of the denominator of the transfer function. The transfer function is

$$\frac{z^{-100}}{1 - 10^{-5} z^{-100}},$$

so its poles are the zeros of $z^{100} - 10^{-5}$. This gives poles

$$z = 10^{-5/100} e^{2\pi k/100}, \quad k = 0, \ldots, 99.$$

These are points uniformly spaced, 3.6° apart, on a circle of radius $10^{-5/100} = 0.8913$. In particular, the eigenvalues are distinct, so A is diagonalizable. Note that the eigenvalues of this A, compared to the case without feedback, are *dramatically different*. Note also that a small change to the matrix (chaging one entry from 0 to 10^{-5}) has a huge change on the eigenvalues.

It's very easy to find the impulse response of the system with feedback. If we put in a unit pulse at $t = 0$, a unit pulse comes out at $t = 100$. Then at $t = 200$, another pulse comes out with amplitude 10^{-5}. At $t = 300$, we get another one, with amplitude 10^{-10}. And so on. In other words, we get a *very* faint echo, with a period of 100. The first echo is barely detectable, and the second and subsequent ones are not present, for any practical purpose. Thus, the feedback has essentially *no effect* on the system.

This impulse response is not what you might have predicted from the eigenvalues. While the magnitude of the eigenvalues makes sense, most of them are complex, so we'd expect to see some strong oscillations. But the impulse response is much simpler than you might have guessed.

We've seen the feedback has essentially no effect on the impulse response of the system, but it causes an extreme change in eigenvalues. Why? There are lots of answers, but the most important message is that the poles can give you a good qualitative feel for the impulse response, when there's just a handful of them. But when there are many of them, it need not.

The extreme sensitivity of the eigenvalues to the change in the one entry is due to the Jordan form of A. When a matrix is diagonalizable, the eigenvalues are differentiable functions of the entries. But when the matrix has Jordan blocks, the eigenvalues can change dramatically with small changes in the matrix.

14.2 *Norm expressions for quadratic forms.* Let $f(x) = x^T Ax$ (with $A = A^T \in \mathbf{R}^{n\times n}$) be a quadratic form.

(a) Show that f is positive semidefinite (*i.e.*, $A \geq 0$) if and only if it can be expressed as $f(x) = \|Fx\|^2$ for some matrix $F \in \mathbf{R}^{k\times n}$. Explain how to find such an F (when $A \geq 0$). What is the size of the smallest such F (*i.e.*, how small can k be)?

(b) Show that f can be expressed as a difference of squared norms, in the form $f(x) = \|Fx\|^2 - \|Gx\|^2$, for some appropriate matrices F and G. How small can the sizes of F and G be?

Solution:

(a) We know that the norm expression $f(x) = \|Fx\|^2$ is a positive semidefinite quadratic form simply because $f(x) \geq 0$ for all x and $f(x) = x^T Ax$ with $A = F^T F \geq 0$. In this problem we will show the converse, *i.e.*, any positive semidefinite quadratic form $f(x) = x^T Ax$ can be written as a norm expression $f(x) = \|Fx\|^2$. Suppose the eigenvalue decomposition of $A \geq 0$ is $Q\Lambda Q^T$, with $Q^T Q = I$ and $\Lambda = \mathbf{diag}(\lambda_1, \ldots, \lambda_n)$ where λ_i are the eigenvalues of A. Since $\lambda_i \geq 0$ (because $A \geq 0$) then $\Lambda^{1/2} = \mathbf{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$ is a real matrix. Let $F = \Lambda^{1/2}Q^T \in \mathbf{R}^{n\times n}$. Then we have $\|Fx\|^2 = x^T F^T Fx = Q\Lambda^{1/2}\Lambda^{1/2}Q^T = x^T Ax = f(x)$. To get smallest F suppose that $\mathbf{Rank}(A) = r$. Therefore, $A \in \mathbf{R}^{n\times n}$ has exactly r nonzero eigenvalues $\lambda_1, \ldots, \lambda_r$. Suppose $\Lambda_+ = \mathbf{diag}(\lambda_1, \ldots, \lambda_r)$. Hence, the eigenvalue decomposition of A can be written as

$$A = \left[\begin{array}{c|c} Q_1 & Q_2 \end{array}\right] \left[\begin{array}{c|c} \Lambda_+ & 0_{r\times(n-r)} \\ \hline 0_{(n-r)\times r} & 0_{(n-r)\times(n-r)} \end{array}\right] \left[\begin{array}{c} Q_1^T \\ \hline Q_2^T \end{array}\right]$$

and as a result $A = Q_1\Lambda_+ Q_1^T$ where $Q_1 \in \mathbf{R}^{n\times r}$. Now we can take $F = \Lambda_+^{1/2}Q_1^T \in \mathbf{R}^{r\times n}$. Therefore, k can be as small as r, *i.e.*, $k = \mathbf{Rank}(r)$. Note that k cannot be any smaller than $\mathbf{Rank}(A)$ because $A = F^T F$ implies that $\mathbf{Rank}(A) \leq k$.

(b) In general, a quadratic form need not to be positive semidefinite. In this problem we show that any quadratic form can be decomposed into its "positive" and "negative" parts. In other words, we can write $f(x)$ as the difference of two norm expressions, *i.e.*, $f(x) = \|Fx\|^2 - \|Gx\|^2$. Suppose A has n_1 positive eigenvalues $\lambda_1, \ldots, \lambda_{n_1}$, n_2 negative eigenvalues $\lambda_{n_1+1}, \ldots, \lambda_{n_1+n_2}$, and therefore $n - n_1 - n_2$ zero eigenvalues. Let

$$\Lambda_+ = \mathbf{diag}(\lambda_1, \ldots, \lambda_{n_1}), \quad \Lambda_- = \mathbf{diag}(-\lambda_{n_1+1}, \ldots, -\lambda_{n_1+n_2}).$$

The eigenvalue decomposition of A can be written as

$$A = \left[\begin{array}{c|c|c} Q_1 & Q_2 & Q_3 \end{array}\right] \left[\begin{array}{c|c|c} \Lambda_+ & 0_{n_1\times n_2} & 0_{n_1\times(n-n_1-n_2)} \\ \hline 0_{n_2\times n_1} & -\Lambda_- & 0_{n_2\times(n-n_1-n_2)} \\ \hline 0_{(n-n_1-n_2)\times n_1} & 0_{(n-n_1-n_2)\times n_2} & 0_{(n-n_1-n_2)\times(n-n_1-n_2)} \end{array}\right] \left[\begin{array}{c} Q_1^T \\ \hline Q_2^T \\ \hline Q_3^T \end{array}\right]$$

so $A = Q_1 \Lambda_+ Q_1^T - Q_2^T \Lambda_- Q_2$. Now simply take $F = \Lambda_+^{1/2} Q_1^T \in \mathbf{R}^{n_1 \times n}$ and $G = \Lambda_-^{1/2} Q_2^T \in \mathbf{R}^{n_2 \times n}$. It is easy to verify that $A = F^T F - G^T G$ and therefore $x^T A x = \|Fx\|^2 - \|Gx\|^2$. In fact, this method gives the smallest sizes for F and G.

14.3 *Congruences and quadratic forms.* Suppose $A = A^T \in \mathbf{R}^{n \times n}$.

(a) Let $Z \in \mathbf{R}^{n \times p}$ be any matrix. Show that $Z^T A Z \geq 0$ if $A \geq 0$.

(b) Suppose that $T \in \mathbf{R}^{n \times n}$ is invertible. Show that $T^T A T \geq 0$ *if and only if* $A \geq 0$. When T is invertible, TAT^T is called a *congruence* of A and TAT^T and A are said to be *congruent*. This problem shows that congruences preserve positive semidefiniteness.

Solution:

(a) By definition, all we have to show is that $x^T (Z^T A Z) x \geq 0$ for all $x \in \mathbf{R}^p$. But $x^T (Z^T A Z) x = (Zx)^T A (Zx)$ and by considering Zx as an element in $\mathbf{R}^n$, since $A \geq 0$ we have $(Zx)^T A (Zx) \geq 0$ and we are done.

(b) The "if" part was shown in problem (0a) (simply take $Z = T$). For the "only if" part we have to show that $T^T A T \geq 0$ implies $A \geq 0$. By definition, it suffices to prove that $x^T A x \geq 0$ for all $x \in \mathbf{R}^n$. Suppose that x is an arbitrary element in $\mathbf{R}^n$. Since T is invertible, T^{-1} exists and let $y = T^{-1} x \in \mathbf{R}^n$. Using the fact that $T^T A T \geq 0$ we have $y^T (T^T A T) y \geq 0$. But

$$y^T (T^T A T) y = (T^{-1} x)^T (T^T A T)(T^{-1} x) = x^T T^{-T} (T^T A T) T^{-1} x = x^T A x,$$

and therefore $y^T (T^T A T) y \geq 0$ implies $x^T A x \geq 0$ and we are done.

14.4 *Positive semidefinite (PSD) matrices.*

(a) Show that if A and B are PSD and $\alpha \in \mathbf{R}$, $\alpha \geq 0$, then so are αA and $A + B$.

(b) Show that any (symmetric) submatrix of a PSD matrix is PSD. (To form a symmetric submatrix, choose any subset of $\{1, \ldots, n\}$ and then throw away all other columns and rows.)

(c) Show that if $A \geq 0$, $A_{ii} \geq 0$.

(d) Show that if $A \geq 0$, $|A_{ij}| \leq \sqrt{A_{ii} A_{jj}}$. In particular, if $A_{ii} = 0$, then the entire ith row and column of A are zero.

Solution:

(a) To show that $\alpha A \geq 0$ we verify that $x^T (\alpha A) x \geq 0$ for all x. But $x^T (\alpha A) x = \alpha (x^T A x)$ and since $x^T A x \geq 0$ ($A \geq 0$) and $\alpha \geq 0$, we immediately get $x^T (\alpha A) x \geq 0$. Again, to show that $A + B \geq 0$ we show that $x^T (A + B) x \geq 0$ for all x. This is easy because $x^T (A + B) x = x^T A x + x^T B x$ and $A, B \geq 0$ imply that $x^T A x, x^T B x \geq 0$ and therefore $x^T (A + B) x \geq 0$.

(b) Suppose that $A = A^T \geq 0$. Any symmetric submatrix of A can be written as $Z^T AZ$ for some suitable matrix Z. For example, if $A \in \mathbf{R}^{3\times 3}$ and we want to pick the submatrix formed by the first and third columns and rows we simply take

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

so that

$$Z^T AZ = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{12} & A_{22} & A_{23} \\ A_{13} & A_{23} & A_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{13} \\ A_{13} & A_{33} \end{bmatrix}.$$

The idea here is to pick the columns of Z as the unit vectors corresponding to the column/row numbers we want to keep. In this example, we wanted to keep the first and third columns/rows so we took $Z = [e_1\ e_3]$. In general, consider the $m \times m$ symmetric submatrix of A which consists of elements of A that are only on the columns and rows $i_1, \ldots, i_m$ of A. Then it is easy to verify that

$$\text{(submatrix formed from columns/rows } i_1, \ldots, i_m) = Z^T AZ, \quad Z = [e_{i_1}\ \cdots\ e_{i_m}],$$

where e_{i_j} is the i_jth unit vector in $\mathbf{R}^n$. Using the result of problem (0a), $A \geq 0$ implies that $Z^T AZ \geq 0$ and therefore any symmetric submatrix of A is also positive semidefinite.

(c) This is easy. We can simply use the result of the previous part ($A_{ii} \in \mathbf{R}$ is a 1×1 symmetric submatrix of A), or more directly, use the fact that $A \geq 0$ implies $e_i^T Ae_i \geq 0$ and note that $e_i^T Ae_i$ is nothing but A_{ii}.

(d) Choose any 2×2 symmetric submatrix of A, say

$$\tilde{A} = \begin{bmatrix} A_{ii} & A_{ij} \\ A_{ij} & A_{jj} \end{bmatrix}.$$

According to problem (0b) this (symmetric) submatrix is positive semidefinite and therefore its eigenvalues are nonnegative. Hence, the determinant of the submatrix (which is equal to the product of the eigenvalues) is also nonnegative. In other words

$$\det \tilde{A} = \begin{bmatrix} A_{ii} & A_{ij} \\ A_{ij} & A_{jj} \end{bmatrix} = A_{ii}A_{jj} - A_{ij}^2 \geq 0$$

and immediately we get $A_{ij}^2 \leq A_{ii}A_{jj}$ or $|A_{ij}| \leq \sqrt{A_{ii}A_{jj}}$. In particular, if $A_{ii} = 0$ then $|A_{ij}| \leq 0$ or $A_{ij} = 0$ (for any j) and the entire ith row (and hence ith column since A is symmetric) should be zero.

14.6 *Gram matrices.* Given functions $f_i : [a,b] \to \mathbf{R}$, $i = 1,\dots,n$, the *Gram matrix* $G \in \mathbf{R}^{n\times n}$ associated with them is defined by

$$G_{ij} = \int_a^b f_i(t) f_j(t)\, dt.$$

(a) Show that $G = G^T \geq 0$.

(b) Show that G is singular if and only if the functions $f_1,\dots,f_n$ are linearly dependent.

Solution:

(a) First of all it is obvious that $G = G^T$ because

$$G_{ij} = \int_a^b f_i(t) f_j(t)\, dt = \int_a^b f_j(t) f_i(t)\, dt = G_{ji}.$$

Define the vector function $f : [a,b] \to \mathbf{R}^n$ as $f(t) = [\, f_1(t) \quad f_2(t) \quad \cdots \quad f_n(t)\,]^T$. Hence

$$G = \int_a^b f(t) f(t)^T\, dt$$

since

$$\begin{aligned} f(t)f(t)^T &= \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix} \begin{bmatrix} f_1(t) & f_2(t) & \cdots & f_n(t) \end{bmatrix} \\ &= \begin{bmatrix} f_1(t)f_1(t) & f_1(t)f_2(t) & \cdots & f_1(t)f_n(t) \\ f_2(t)f_1(t) & f_2(t)f_2(t) & \cdots & f_2(t)f_n(t) \\ \vdots & & & \vdots \\ f_n(t)f_1(t) & f_n(t)f_2(t) & \cdots & f_n(t)f_n(t) \end{bmatrix}. \end{aligned}$$

Now to show that $G \geq 0$, we consider any $x = [\, x_1 \quad \cdots \quad x_n\,]^T \in \mathbf{R}^n$ and verify that $x^T G x \geq 0$. We have

$$x^T G x = x^T \Big(\int_a^b f(t) f(t)^T\, dt\Big) x = \int_a^b x^T f(t) f(t)^T x\, dt = \int_a^b (x^T f(t))^2\, dt,$$

and therefore $x^T G x \geq 0$ since it is the area under the nonnegative function $(x^T f(t))^2$ from a to b.

(b) Note that if $G \geq 0$ is nonsingular it does not have any zero eigenvalues (in fact all eigenvalues should be strictly positive) and therefore $x^T G x > 0$ for all x or $G > 0$. We show that if $f_1,\dots,f_n$ are linearly independent then $G > 0$ and vice versa. In the previous part we showed that

$$x^T G x = \int_a^b (x^T f(t))^2\, dt = \int_a^b (x_1 f_1(t) + \cdots + x_n f_n(t))^2\, dt.$$

Therefore, $x^TGx > 0$ for all x as long as no linear combination of the functions $f_1, \dots, f_n$ is zero, *i.e.*, $x_1f_1(t) + \cdots + x_nf_n(t) \neq 0$. In other words, $G > 0$ if and only if the functions $f_1, \dots, f_n$ are linearly independent and we are done.

14.8 Express $\sum_{i=1}^{n-1}(x_{i+1}-x_i)^2$ in the form x^TPx with $P = P^T$. Is $P \geq 0$? $P > 0$? *Solution:* Let $x = [x_1 \ \cdots \ x_n]^T \in \mathbf{R}^n$. We have

$$\begin{aligned}
\sum_{i=1}^{n-1}(x_{i+1}-x_i)^2 &= \sum_{i=1}^{n-1}(e_{i+1}^Tx - e_i^Tx)^2 \\
&= \sum_{i=1}^{n-1}\left((e_{i+1}^T - e_i^T)x\right)^2 \\
&= \sum_{i=1}^{n-1}\left((e_{i+1}^T - e_i^T)x\right)^T\left((e_{i+1}^T - e_i^T)x\right) \\
&= \sum_{i=1}^{n-1} x^T(e_{i+1} - e_i)(e_{i+1}^T - e_i^T)x \\
&= x^T\left(\sum_{i=1}^{n-1}(e_{i+1} - e_i)(e_{i+1}^T - e_i^T)\right)x \\
&= x^T\left(\sum_{i=1}^{n-1}(e_{i+1}e_{i+1}^T - e_{i+1}e_i^T - e_ie_{i+1}^T + e_ie_i^T)\right)x \\
&= x^TPx.
\end{aligned}$$

Let $E(i,j) := e_ie_j^T \in \mathbf{R}^{n\times n}$ (*i.e.*, a matrix with all zero elements except for a "1" at the (i,j)th entry). The matrix $P \in \mathbf{R}^{n\times n}$ can be written in terms of $E(i,j)$ as follows:

$$\begin{aligned}
P &= \sum_{i=1}^{n-1}(e_{i+1}e_{i+1}^T - e_{i+1}e_i^T - e_ie_{i+1}^T + e_ie_i^T) \\
&= \sum_{i=1}^{n-1}\left(E(i+1,i+1) - E(i+1,i) - E(i,i+1) + E(i,i)\right) \\
&= \sum_{i=1}^{n-1}E(i+1,i+1) - \sum_{i=1}^{n-1}E(i+1,i) - \sum_{i=1}^{n-1}E(i,i+1) + \sum_{i=1}^{n-1}E(i,i).
\end{aligned}$$

But

$$\sum_{i=1}^{n-1}E(i+1,i+1) = \begin{bmatrix} 0 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}, \quad \sum_{i=1}^{n-1}E(i,i) = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 0 \end{bmatrix}$$

and

$$\sum_{i=1}^{n-1} E(i+1,i) = \begin{bmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & & \ddots & \\ & & & 1 & 0 \end{bmatrix}, \quad \sum_{i=1}^{n-1} E(i,i+1) = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}.$$

Therefore

$$P = \begin{bmatrix} 1 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{bmatrix}$$

is a tridiagonal matrix. Clearly, since the quadratic form is nothing but the sum of squares ($\sum_{i=1}^{n-1}(x_{i+1}-x_i)^2$) it is nonnegative for any x and therefore $P \geq 0$. However, P is *not* strictly greater than zero because for example $x = [1 \ \cdots \ 1]^T \neq 0$ gives $x^T P x = \sum_{i=1}^{n-1}(1-1)^2 = 0$.

14.9 Suppose A and B are symmetric matrices that yield the same quadratic form, *i.e.*, $x^T A x = x^T B x$ for all x. Show that $A = B$. *Hint:* first try $x = e_i$ (the ith unit vector) to conclude that the entries of A and B on the diagonal are the same; then try $x = e_i + e_j$. *Solution:*
With $x = e_i$, $x^T A x = x^T B x$ gives

$$A_{ii} = B_{ii}, \quad i = 1, 2, ..., n.$$

With $x = e_i + e_j$, $x^T A x = x^T B x$ gives

$$e_i^T A e_i + e_i^T A e_j + e_j^T A e_i + e_j^T A e_j = e_i^T B e_i + e_i^T B e_j + e_j^T B e_i + e_j^T B e_j$$

and therefore

$$A_{ii} + A_{ij} + A_{ji} + A_{jj} = B_{ii} + B_{ij} + B_{ji} + B_{jj}, \quad i, j = 1, 2, ..., n,$$

so that $A_{ij} + A_{ji} = B_{ij} + B_{ji}$. This, along with $A_{ii} = B_{ii}$ means that $A + A^T = B + B^T$. Finally, using the fact that $A = A^T$ and $B = B^T$, we conclude that $A = B$.

14.11 Suppose that $A \in \mathbf{R}^{n \times n}$. Show that $\|A\| < 1$ implies $I + A$ is invertible. *Solution:*
We show that $\|A\| < 1$ implies $I+A$ invertible by proving the contrapositive statement: $I + A$ singular implies $\|A\| \geq 1$. The proof is very simple this way. For $I + A$ singular, there exists a nonzero vector v such that $(I + A)v = 0$, or $Av = -v$. For this v, we have $\frac{\|Av\|}{\|v\|} = 1$. Now, $\frac{\|Av\|}{\|v\|} \leq \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \|A\|$. Thus, $\|A\| \geq 1$, which proves the contrapositive.

14.13 *Eigenvalues and singular values of a symmetric matrix.* Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues, and let $\sigma_1, \ldots, \sigma_n$ be the singular values of a matrix $A \in \mathbf{R}^{n\times n}$, which satisfies $A = A^T$. (The singular values are based on the full SVD: If $\mathbf{Rank}(A) < n$, then some of the singular values are zero.) You can assume the eigenvalues (and of course singular values) are sorted, *i.e.*, $\lambda_1 \geq \cdots \geq \lambda_n$ and $\sigma_1 \geq \cdots \geq \sigma_n$. How are the eigenvalues and singular values related? *Solution:*
Since A is symmetric it can be diagonalized using an orthogonal matrix Q as

$$A = Q \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} Q^T$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A. Suppose that the columns of Q are ordered such that $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$. Thus

$$A = Q \begin{bmatrix} \mathrm{sgn}\lambda_1 & & \\ & \ddots & \\ & & \mathrm{sgn}\lambda_n \end{bmatrix} \begin{bmatrix} |\lambda_1| & & \\ & \ddots & \\ & & |\lambda_n| \end{bmatrix} Q^T$$

Now we define

$$U = Q \begin{bmatrix} \mathrm{sgn}\lambda_1 & & \\ & \ddots & \\ & & \mathrm{sgn}\lambda_n \end{bmatrix}, \quad \Sigma = \begin{bmatrix} |\lambda_1| & & \\ & \ddots & \\ & & |\lambda_n| \end{bmatrix}, \quad V = Q.$$

Clearly, U is a an orthogonal matrix because $UU^T = QQ^T = I$. Now $A = U\Sigma V^T$is a SVD of A, and we conclude that $\sigma_i = |\lambda_i|$.

14.21 *Two representations of an ellipsoid.* In the lectures, we saw two different ways of representing an ellipsoid, centered at 0, with non-zero volume. The first uses a quadratic form:

$$\mathcal{E}_1 = \left\{ x \,\middle|\, x^T S x \leq 1 \right\},$$

with $S^T = S > 0$. The second is as the image of a unit ball under a linear mapping:

$$\mathcal{E}_2 = \{ y \mid y = Ax, \|x\| \leq 1 \},$$

with $\det(A) \neq 0$.

(a) Given S, explain how to find an A so that $\mathcal{E}_1 = \mathcal{E}_2$.

(b) Given A, explain how to find an S so that $\mathcal{E}_1 = \mathcal{E}_2$.

(c) What about uniqueness? Given S, explain how to find *all* A that yield $\mathcal{E}_1 = \mathcal{E}_2$. Given A, explain how to find *all* S that yield $\mathcal{E}_1 = \mathcal{E}_2$.

Solution:
First we will show that

$$\mathcal{E}_2 = \left\{ y \mid y^T(AA^T)^{-1}y \le 1 \right\} \tag{1}$$

$$\begin{aligned}
\mathcal{E}_2 &= \{y \mid y = Ax, \|x\| \le 1\} \\
&= \left\{ y \mid A^{-1}y = x, \|x\| \le 1 \right\} \text{ since } A \text{ is invertible square matrix} \\
&= \left\{ y \mid \|A^{-1}y\| \le 1 \right\} \\
&= \left\{ y \mid y^T A^{-T} A^{-1} y \le 1 \right\} = \left\{ y \mid y^T (AA^T)^{-1} y \le 1 \right\}
\end{aligned}$$

Since S is symmetric positive definite, the eigenvalues of S are all positive and we can choose n orthonormal eigenvectors. So $S = Q\Lambda Q^T$ where $\Lambda = \mathbf{diag}(\lambda_1, \ldots, \lambda_n) > 0$ and Q is orthogonal. Let $\Lambda^{\frac{1}{2}} = \mathbf{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$. If we let $A = Q(\Lambda^{\frac{1}{2}})^{-1} = Q\Lambda^{-\frac{1}{2}}$,

$$\begin{aligned}
(AA^T)^{-1} &= (Q\Lambda^{-\frac{1}{2}}\Lambda^{-\frac{1}{2}}Q^T)^{-1} \\
&= (Q\Lambda^{-1}Q^T)^{-1} = Q\Lambda Q^T \\
&= S
\end{aligned}$$

Therefore, by (1) $\mathbf{A} = \mathbf{Q}\boldsymbol{\Lambda}^{-\frac{1}{2}}$ yields $\mathcal{E}_1 = \mathcal{E}_2$. By (1), $\mathbf{S} = (\mathbf{AA^T})^{-\mathbf{1}}$ yields $\mathcal{E}_1 = \mathcal{E}_2$. *Uniqueness:* We show that

$$\mathcal{E}_S = \mathcal{E}_T \Leftrightarrow S = T \tag{2}$$

where $\mathcal{E}_S = \{x \mid x^T S x \le 1\}$, $\mathcal{E}_T = \{x \mid x^T T x \le 1\}$, $S^T = S > 0$ and $T^T = T > 0$. It is clear that if $S = T$, then $\mathcal{E}_S = \mathcal{E}_T$. Now we show that $\mathcal{E}_S = \mathcal{E}_T \Rightarrow x^T S x = x^T T x$, $\forall x \in \mathbf{R}^n$. Without loss of generality let's assume $\exists x_0 \in \mathbf{R}^n$ such that $x_0^T S x_0 > x_0^T T x_0 = \alpha \neq 0$. If we let $x_1 = x_0/\sqrt{\alpha}$, then $x_1^T T x_1 = 1$, but $x_1^T S x_1 > 1$, thus $x_1 \in \mathcal{E}_T$ but $x_1 \notin \mathcal{E}_S$, and therefore $\mathcal{E}_S \neq \mathcal{E}_T$. Finally, $\mathcal{E}_S = \mathcal{E}_T \Rightarrow x^T S x = x^T T x$, $\forall x \in \mathbf{R}^n \Rightarrow S = T$ by the uniqueness of the symmetric part in a quadratic form. Hence S is unique. *Given S, find all A that yield $\mathcal{E}_1 = \mathcal{E}_2$.*
The answer is

$$A = Q\Lambda^{-\frac{1}{2}}V^T$$

where $V \in \mathbf{R}^{n\times n}$ is any orthogonal matrix and

$$S^T = S = Q\Lambda Q^T > 0$$

where Q is orthogonal and $\Lambda = \mathbf{diag}(\lambda_1, \ldots, \lambda_n) > 0$. Let the singular value decomposition of A be

$$A = U\Sigma V^T$$

where $U, V \in \mathbf{R}^{n\times n}$ are orthogonal and $\Sigma = \mathbf{diag}(\sigma_1, \ldots, \sigma_n) > 0$ (since $\det(A) \neq 0$.) By (1),

$$\begin{aligned}
\mathcal{E}_2 &= \left\{ y \mid y^T(AA^T)^{-1}y \le 1 \right\} \\
&= \left\{ y \mid y^T(U\Sigma^2U^T)^{-1}y \le 1 \right\} \\
&= \left\{ y \mid y^T U\Sigma^{-2}U^T y \le 1 \right\}
\end{aligned}$$

Thus, if $\mathcal{E}_1 = \mathcal{E}_2$, then $S = U\Sigma^{-2}U^T$ by (2). Therefore $U = Q$ and $\Sigma = \Lambda^{-\frac{1}{2}}$, and V can be any orthogonal matrix. You can also see why A's are different only by right-side multiplication by an orthogonal matrix by the following argument. By (2) and (1), $AA^T = S^{-1}$. Let

$$A = \left[\begin{array}{cccc} \tilde{a}_1 & \tilde{a}_2 & \ldots & \tilde{a}_n \end{array} \right]^T$$

Then we have,

$$\begin{aligned} \|\tilde{a}_i\|^2 &= (S^{-1})_{ii}, \\ \tilde{a}_i^T \tilde{a}_j &= (S^{-1})_{ij}, \end{aligned}$$

and

$$\cos\theta_{ij} = \frac{(S^{-1})_{ij}}{\sqrt{(S^{-1})_{ii}(S^{-1})_{jj}}}.$$

This means that the row vectors of any A satisfying $AA^T = S^{-1}$ have the same length and the same angle between any two of them. So the rows of A can vary only by the application of an identical rotation or reflection to all of them. These are the transformations preserving length and angle, and correspond to orthogonal matrices. Since we are considering row vectors, the orthogonal matrix should be multiplied on the right.

14.33 *Uncovering a hidden linear explanation.* Consider a set of vectors $y_1, \ldots, y_N \in \mathbf{R}^n$, which might represent a collection of measurements or other data. Suppose we have

$$y_i \approx Ax_i + b, \quad i = 1, \ldots, N,$$

where $A \in \mathbf{R}^{n\times m}$, $x_i \in \mathbf{R}^m$, and $b \in \mathbf{R}^n$, with $m < n$. (Our main interest is in the case when N is much larger than n, and m is smaller than n.) Then we say that $y = Ax + b$ is a *linear explanation* of the data y. We refer to x as the vector of *factors* or *underlying causes* of the data y. For example, suppose $N = 500$, $n = 30$, and $m = 5$. In this case we have 500 vectors; each vector y_i consists of 30 scalar measurements or data points. But these 30-dimensional data points can be 'explained' by a much smaller set of 5 'factors' (the components of x_i). This problem is about uncovering, or discovering, a linear explanation of a set of data, given only the data. In other words, we are given $y_1, \ldots, y_N$, and the goal is to find m, A, b, and $x_1, \ldots, x_N$ so that $y_i \approx Ax_i + b$. To judge the accuracy of a proposed explanation, we'll use the RMS explanation error, *i.e.*,

$$J = \left(\frac{1}{N} \sum_{i=1}^{N} \|y_i - Ax_i - b\|^2 \right)^{1/2}.$$

One rather simple linear explanation of the data is obtained with $x_i = y_i$, $A = I$, and $b = 0$. In other words, the data explains itself! In this case, of course, we have $y_i = Ax_i + b$, so the RMS explanation error is zero. But this is not what we're after. To be a useful explanation, we want to have m substantially smaller than n, *i.e.*, substantially fewer factors than the dimension of the original data (and for this smaller dimension, we'll accept a nonzero, but hopefully small, value of J.) Generally, we want m, the number of factors in the explanation, to be as small as possible, subject to the constraint that J is not too large. Even if we fix the number of factors as m, a linear explanation of a set of data is not unique. Suppose A, b, and $x_1, \dots, x_N$ is a linear explanation of our data, with $x_i \in \mathbf{R}^m$. Then we can multiply the matrix A by two (say), and divide each vector x_i by two, and we have another linear explanation of the original data. More generally, let $F \in \mathbf{R}^{m \times m}$ be invertible, and $g \in \mathbf{R}^m$. Then we have

$$y_i \approx Ax_i + b = (AF^{-1})(Fx_i + g) + (b - AF^{-1}g).$$

Thus,

$$\tilde{A} = AF^{-1}, \qquad \tilde{b} = b - AF^{-1}g, \qquad \tilde{x}_1 = Fx_1 + g, \quad \dots, \quad \tilde{x}_N = Fx_N + g$$

is another equally good linear explanation of the data. In other words, we can apply any affine (*i.e.*, linear plus constant) mapping to the underlying factors x_i, and generate another equally good explanation of the original data by appropriately adjusting A and b. To standardize or normalize the linear explanation, it is usually assumed that

$$\frac{1}{N}\sum_{i=1}^{N} x_i = 0, \qquad \frac{1}{N}\sum_{i=1}^{N} x_i x_i^T = I.$$

In other words, the underlying factors have an average value zero, and unit sample covariance. (You don't need to know what covariance is — it's just a definition here.) Finally, the problem.

(a) Explain clearly how you would find a hidden linear explanation for a set of data $y_1, \dots, y_N$. Be sure to say how you find m, the dimension of the underlying causes, the matrix A, the vector b, and the vectors $x_1, \dots, x_N$. Explain clearly why the vectors $x_1, \dots, x_N$ have the required properties.

(b) Carry out your method on the data in the file `linexp_data.m` available on the course web site. The file gives the matrix $Y = [y_1 \ \cdots \ y_N]$. Give your final A, b, and $x_1, \dots, x_N$, and verify that $y_i \approx Ax_i - b$ by calculating the norm of the error vector, $\|y_i - Ax_i - b\|$, for $i = 1, \dots, N$. Sort these norms in descending order and plot them. (This gives a good picture of the distribution of explanation errors.) By explicit computation verify that the vectors $x_1, \dots, x_N$ obtained, have the required properties.

Solution.

(a) We have to find b and $x_1, \dots, x_N$ that minimize

$$\tilde{J} = NJ^2 = \sum_{i=1}^{N}(y_i - Ax_i - b)^T(y_i - Ax_i - b).$$

Taking the gradient with respect to b and setting it to zero gives,

$$\nabla_b(\tilde{J}) = \sum_{i=1}^{N} 2(y_i - Ax_i - b)(-1) = 0.$$

Since we want $\sum_{i=1}^{N} x_i = 0$, we get

$$b = \frac{1}{N}\sum_{i=1}^{N} y_i.$$

Let X be the matrix $[x_1 \ \cdots \ x_N] \in \mathbf{R}^{m\times N}$. Let $z_i = y_i - b,\ i = 1,\ldots,N$, and $Z = \left[\begin{array}{ccc} z_1 & \cdots & z_N \end{array}\right] \in \mathbf{R}^{n\times N}$. Note that Z is known from the data after b is calculated as shown above. The matrix $(AX) \in \mathbf{R}^{n\times N}$ and has rank at most rank m. Then

$$\tilde{J} = \sum_{i=1}^{N} \|z_i - Ax_i\|^2 = \sum_{k=1}^{n}\sum_{j=1}^{N}(Z_{kj} - (AX)_{kj})^2 = \|Z - AX\|_F^2.$$

Thus minimizing $\tilde{J}$ is minimizing the Frobenius norm of the matrix $(Z - AX)$ where AX is at most rank m. Let the SVD of Z be $Z = U\Sigma V^T$, where $U \in \mathbf{R}^{n\times r}$, $\Sigma \in \mathbf{R}^{r\times r}$, $V \in \mathbf{R}^{N\times r}$ and r is the rank of Z. The choice of m depends on the singular values $\sigma_1,\ldots,\sigma_r$ obtained for the particular data. A good choice of m would be when there is a significant jump in the singular values, *i.e.*, $\sigma_m \gg \sigma_{m+1}$; or when the singular value becomes small enough (σ_{m+1} is negligible). Thus we pick a value for m. Then the m rank approximation to Z is

$$(AX) = \sum_{i=1}^{m} \sigma_i u_i v_i^T = U_m\Sigma_m V_m^T,$$

where $U_m \in \mathbf{R}^{n\times m}$, $\Sigma_m \in \mathbf{R}^{m\times m}$, $V_m \in \mathbf{R}^{N\times m}$. We pick $A = \frac{1}{\sqrt{N}}U_m\Sigma_m$, and x_i as $\sqrt{N}$ times the ith column of V_m^T. Then

$$\frac{1}{N}\sum_{i=1}^{N} x_i x_i^T = \frac{1}{N}(\sqrt{N}V_m)^T(\sqrt{N}V_m) = I.$$

In order to show that $\frac{1}{N}\sum_{i=1}^{N} x_i = 0$, consider

$$Z\mathbf{1} = \sum_{i=1}^{N} z_i = \sum_{i=1}^{N}(y_i - b) = 0,$$

where $\mathbf{1}$ is the vector of ones of size N. The vector $\mathbf{1}$ is in the nullspace of Z which we can write as $U\Sigma V^T\mathbf{1} = 0$. The matrix $U\Sigma$ is full rank, therefore $V^T\mathbf{1} = 0$. Hence $V_m^T\mathbf{1} = 0$ as V_m are the first m columns of the matrix V. This means

$$\frac{1}{\sqrt{N}}V_m^T\mathbf{1} = \frac{1}{N}\left[x_1 \ \cdots \ x_N\right]\mathbf{1} = \frac{1}{N}\sum_{i=1}^{N} x_i = 0.$$

Thus we have found

$$b = \frac{1}{N}\sum_{i=1}^{N} y_i, \quad A = \frac{1}{\sqrt{N}} U_m \Sigma_m, \quad [x_1 \ \cdots \ x_N] = \sqrt{N} V_m^T,$$

with the required properties.

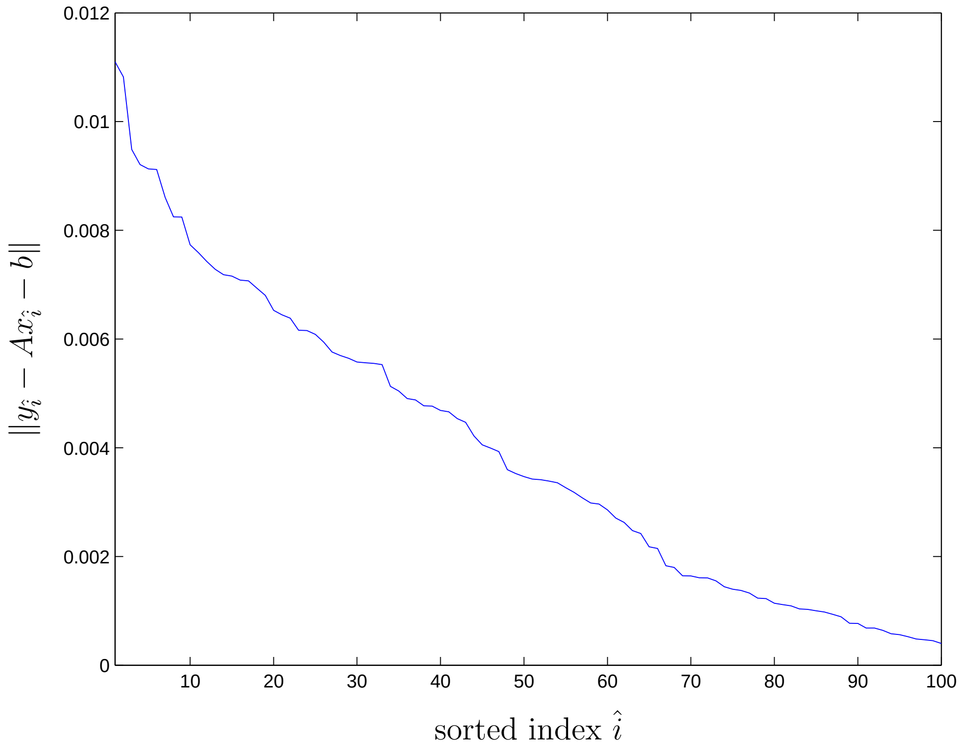
(b) The following Matlab code implements solution method described in the part (a). We observe that there are 3 significant singular values, and therefore we take $m = 3$.

```
linexp_data;
[n,N] = size(Y);
b = sum(Y')'/N;
for i = 1:N
Z(:, i) = Y(:, i) - b;
end
[U S V] = svd(Z);
A = 1/sqrt(N)*U(:, 1:3)*S(1:3, 1:3);
X = sqrt(N)*V(:, 1:3)';
error = Z - A*X;
for i=1:N
errNorm(i) = norm(error(:, i));
end
plot(sort(errNorm,2,'descend'));
xlabel('index'); ylabel('errors');
```

Here is explicit check shows that X satisfies given properties.

```
>> sum(X,2)/N =
1.0e-15 *
-0.0333
-0.1064
0.2331
>> X*X'/N =
1.0000    0.0000   -0.0000
0.0000    1.0000    0.0000
-0.0000    0.0000    1.0000
```

Plot of the sorted error norm is shown below:



Solution to additional exercise

1. *Some simple matrix inequality counter-examples.*

   (a) Find a (square) matrix A, which has all eigenvalues real and positive, but there is a vector x for which $x^T Ax < 0$. (Give A and x, and the eigenvalues of A.)

   Moral: You cannot use positivity of the eigenvalues of A as a test for whether $x^T Ax \geq 0$ holds for all x.

   What is the correct way to check whether $x^T Ax \geq 0$ holds for all x? (You are allowed to find eigenvalues in this process.)

   (b) Find symmetric matrices A and B for which neither $A \geq B$ nor $B \geq A$ holds.

   Of course, we'd like the simplest examples in each case.

   Solution.

   (a) $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$.

   (b) $A = \mathbf{diag}(1,2)$, $B = \mathbf{diag}(2,1)$.