b. $(A^\dagger)^\dagger = A$.

The above two properties are similar to those that are satisfied by the usual matrix inverse. However, we point out that the property $(A_1A_2)^\dagger = A_2^\dagger A_1^\dagger$ does not hold in general (see Exercise 12.26).

Finally, it is important to note that we can define the generalized inverse in an alternative way, following the definition of Penrose. Specifically, the Penrose definition of the generalized inverse of a matrix $A \in \mathbb{R}^{m\times n}$ is the unique matrix $A^\dagger \in \mathbb{R}^{n\times m}$ satisfying the following properties:

1. $AA^\dagger A = A$;
2. $A^\dagger AA^\dagger = A^\dagger$;
3. $(AA^\dagger)^T = AA^\dagger$;
4. $(A^\dagger A)^T = A^\dagger A$.

The Penrose definition above is equivalent to Definition 12.1 (see Exercise 12.25). For more information on generalized inverses and their applications, we refer the reader to the books by Ben-Israel and Greville [4], and Campbell and Meyer [18].

EXERCISES

12.1 A rock is accelerated to 3, 5, and 6 m/s^2 by applying forces of 1, 2, and 3 N, respectively. Assuming Newton's law $F = ma$, where F is the force and a is the acceleration, estimate the mass m of the rock using the least squares method.

12.2 A spring is stretched to lengths $L = 3, 4$, and 5 cm under applied forces $F = 1$, 2, and 4 N, respectively. Assuming Hooke's law $L = a + bF$, estimate the normal length a and spring constant b using the least squares approach.

12.3 Suppose that we perform an experiment to calculate the gravitational constant g as follows. We drop a ball from a certain height, and measure its distance from the original point at certain time instants. The results of the experiment are shown in the following table.

Time (seconds)	1.00	2.00	3.00
Distance (meters)	5.00	19.5	44.0

The equation relating the distance s and the time t at which s is measured is given by

$$s = \frac{1}{2}gt^2.$$

a. Find a least-squares estimate of g using the experimental results from the abov[e] table.

b. Suppose that we take an additional measurement at time 4.00, and obtain[...] distance of 78.5. Use the recursive least-squares algorithm to calculate [...] updated least-squares estimate of g.

12.4 Suppose we wish to estimate the value of the resistance R of a resistor. Ohn[...] Law states that if V is the voltage across the resistor, and I is the current throu[...] the resistor, then $V = IR$. To estimate R, we apply a 1 amp current through t[...] resistor and measure the voltage across it. We perform the experiment on n volta[...] measuring devices, and obtain measurements of $V_1, \ldots, V_n$. Show that the le[...] squares estimate of R is simply the average of $V_1, \ldots, V_n$.

12.5 The table below shows the stock prices for three companies, X, Y, and [...] tabulated over three days:

	Day		
	1	2	3
X	6	4	5
Y	1	1	3
Z	2	1	2

Suppose an investment analyst proposes a model for the predicting the stock price [...] X based on those of Y and Z:

$$p_X = ap_Y + bp_Z,$$

where p_X, p_Y, and p_Z are the stock prices of X, Y, and Z, respectively, and a, b a[...] real-valued parameters. Calculate the least squares estimate of the parameters a a[...] b based on the data in the above table.

12.6 We are given two mixtures, A and B. Mixture A contains 30% gold, 40% silve[...] and 30% platinum, whereas mixture B contains 10% gold, 20% silver, and 70[...] platinum (all percentages of weight). We wish to determine the ratio of the weig[...] of mixture A to the weight of mixture B such that we have as close as possible to[...] total of 5 ounces of gold, 3 ounces of silver, and 4 ounces of platinum. Formula[...] and solve the problem using the linear least-squares method.

12.7 *Background:* If $Ax + w = b$, where w is a "white noise" vector, then defin[...] the *least-squares estimate* of x given b to be the solution to the problem

$$\text{minimize}_{x} \; \|Ax - b\|^2.$$

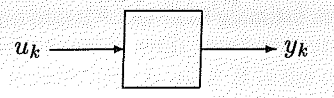

Figure 12.5 Input-output system in Exercise 12.12

a. Suppose we wish to find the best linear estimate of the system based on the above input-output data. In other words, we wish to find a $\hat{\theta}_n \in \mathbb{R}$ to fit the model $y_k = \hat{\theta}_n u_k$, $k = 1, \ldots, n$. Using the least squares approach, derive a formula for $\hat{\theta}_n$ based on $u_1, \ldots, u_n$ and $y_1, \ldots, y_n$.

b. Suppose the data in part a is generated by

$$y_k = \theta u_k + e_k,$$

where $\theta \in \mathbb{R}$ and $u_k = 1$ for all k. Show that the parameter $\hat{\theta}_n$ in part a converges to θ as $n \to \infty$ if and only if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} e_k = 0.$$

12.13 Consider a discrete-time linear system $x_{k+1} = ax_k + bu_k$, where u_k is the input at time k, x_k is the output at time k, and $a, b \in \mathbb{R}$ are system parameters. Suppose that we apply a constant input $u_k = 1$ for all $k \geq 0$, and measure the first 4 values of the output to be $x_0 = 0$, $x_1 = 1$, $x_2 = 2$, $x_3 = 8$. Find the least-squares estimate of a and b based on the above data.

12.14 Consider a discrete-time linear system $x_{k+1} = ax_k + bu_k$, where u_k is the input at time k, x_k is the output at time k, and $a, b \in \mathbb{R}$ are system parameters. Given the first $n+1$ values of the impulse response $h_0, \ldots, h_n$, find the least squares estimate of a and b. You may assume that at least one h_k is nonzero.
Note: The *impulse response* is the output sequence resulting from an input of $u_0 = 1$, $u_k = 0$ for $k \neq 0$, and zero initial condition $x_0 = 0$.

12.15 Consider a discrete-time linear system $x_{k+1} = ax_k + bu_k$, where u_k is the input at time k, x_k is the output at time k, and $a, b \in \mathbb{R}$ are system parameters. Given the first $n+1$ values of the step response $s_0, \ldots, s_n$, where $n > 1$, find the least squares estimate of a and b. You may assume that at least one s_k is nonzero.
Note: The *step response* is the output sequence resulting from an input of $u_k = 1$ for $k \geq 0$, and zero initial condition $x_0 = 0$ (i.e., $s_0 = x_0 = 0$).

12.16 Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $m \leq n$, rank $\boldsymbol{A} = m$, and $\boldsymbol{x}_0 \in \mathbb{R}^n$. Consider the problem

$$\begin{array}{ll} \text{minimize} & \|\boldsymbol{x} - \boldsymbol{x}_0\| \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}. \end{array}$$

Show that the above problem has a unique solution given by

$$\boldsymbol{x}^* = \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T)^{-1}\boldsymbol{b} + (\boldsymbol{I}_n - \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T)^{-1}\boldsymbol{A})\boldsymbol{x}_0.$$

12.17 Given $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $m \leq n$, rank $\boldsymbol{A} = m$, and $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_p \in \mathbb{R}^m$, consider the problem

$$\text{minimize } \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}_1\|^2 + \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}_2\|^2 + \cdots + \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}_p\|^2. \tag{12.1}$$

Suppose that $\boldsymbol{x}_i^*$ is the solution to the problem

$$\text{minimize } \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}_i\|^2,$$

where $i = 1, \ldots, p$. Write the solution to (12.1) in terms of $\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_p^*$.

12.18 Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $m \leq n$, and rank $\boldsymbol{A} = m$. Show that $\boldsymbol{x}^* = \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T)^{-1}\boldsymbol{b}$ is the only vector in $\mathcal{R}(\boldsymbol{A}^T)$ satisfying $\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{b}$.

12.19 Show that in Kaczmarz's algorithm, if $\boldsymbol{x}^{(0)} = \boldsymbol{0}$, then $\boldsymbol{x}^{(k)} \in \mathcal{R}(\boldsymbol{A}^T)$ for all k.

12.20 Consider Kaczmarz's algorithm with $\boldsymbol{x}^{(0)} \neq \boldsymbol{0}$.

a. Show that there exists a unique point minimizing $\|\boldsymbol{x} - \boldsymbol{x}^{(0)}\|$ subject to $\{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$.

b. Show that Kaczmarz's algorithm converges to the point in part a.

12.21 Consider Kaczmarz's algorithm with $\boldsymbol{x}^{(0)} = \boldsymbol{0}$, where $m = 1$; that is, $\boldsymbol{A} = [\boldsymbol{a}^T] \in \mathbb{R}^{1\times n}$ and $\boldsymbol{a} \neq \boldsymbol{0}$, and $0 < \mu < 2$. Show that there exists $0 \leq \gamma < 1$ such that $\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^*\| \leq \gamma\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^*\|$ for all $k \geq 0$.

12.22 Show that in Kaczmarz's algorithm, if $\mu = 1$, then $b_{R(k)+1} - \boldsymbol{a}_{R(k)+1}^T\boldsymbol{x}^{(k+1)} = 0$ for each k.

12.23 Consider the problem of minimizing $\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|^2$ over $\mathbb{R}^n$, where $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$. Let $\boldsymbol{x}^*$ be a solution. Suppose that $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}$ is a full-rank factorization of $\boldsymbol{A}$; that is, rank $\boldsymbol{A}$ = rank $\boldsymbol{B}$ = rank $\boldsymbol{C} = r$, and $\boldsymbol{B} \in \mathbb{R}^{m\times r}$, $\boldsymbol{C} \in \mathbb{R}^{r\times n}$. Show that the minimizer of $\|\boldsymbol{B}\boldsymbol{y} - \boldsymbol{b}\|$ over $\mathbb{R}^r$ is $\boldsymbol{C}\boldsymbol{x}^*$.

12.24 Prove the following properties of generalized inverses:

a. $(\boldsymbol{A}^T)^\dagger = (\boldsymbol{A}^\dagger)^T$;

b. $(\boldsymbol{A}^\dagger)^\dagger = \boldsymbol{A}$.

12.25 Show that the Penrose definition of the generalized inverse is equivalent to Definition 12.1.

12.26 Construct matrices $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ such that $(\boldsymbol{A}_1\boldsymbol{A}_2)^\dagger \neq \boldsymbol{A}_2^\dagger\boldsymbol{A}_1^\dagger$.