Section 6.1

Exercise 10: Find a unit vector $u$ in the direction of the given vector

$$w = \begin{bmatrix} -6 \\ 4 \\ -3 \end{bmatrix}.$$ 

Solution. There are two solutions:

$$u = \pm \frac{1}{\|w\|} w = \pm \frac{1}{\sqrt{36 + 16 + 9}} \begin{bmatrix} -6 \\ 4 \\ -3 \end{bmatrix} = \pm \begin{bmatrix} -6/\sqrt{61} \\ 4/\sqrt{61} \\ -3/\sqrt{61} \end{bmatrix}.$$ \[\hfill \blacksquare\]

Exercise 24: Verify the parallelogram law

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$ 

Solution. Let $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$. We have

$$\|u + v\|^2 + \|u - v\|^2 = \sum_{i=1}^{n} (u_i + v_i)^2 + \sum_{i=1}^{n} (u_i - v_i)^2 = \sum_{i=1}^{n} (2u_i^2 + 2v_i^2) = 2\|u\|^2 + 2\|v\|^2.$$ 

(Here we used the simple identity $(a+b)^2 + (a-b)^2 = 2a^2 + 2b^2$ valid for any scalars $a, b$.) \[\hfill \blacksquare\]

Section 6.2

Exercise 30: Let $U$ be orthonormal matrix, and construct $V$ by interchanging some of the rows of $U$. Explain why $V$ is orthonormal.

Solution 1. Let us recall some implications proved in the class. By definition, a matrix $U$ is orthogonal if and only if $U^T U = I$. Since both $U$ and $U^T$ are square matrices, the latter identity is equivalent to $UU^T = I$ by the Invertible Matrix Theorem. But $UU^T = I$ is equivalent to the rows of $U$ being orthonormal.

In summary, $U$ is orthogonal if and only if its rows are orthonormal. The latter property is clearly preserved by any row permutation.

Solution 2 (or rather a hint). Observe that the dot product $x \cdot y$ does not change if the entries of $x$ are permuted in the same way as the entries of $y$. \[\hfill \blacksquare\]
Section 6.3

Exercise 7: Let \( W = \text{Span}\{u_2, u_2\} \). Write \( y \) as the sum of a vector in \( W \) and a vector orthogonal to \( W \). Here

\[
y = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}, \quad u_1 = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}, \quad u_2 = \begin{bmatrix} 5 \\ 1 \\ 4 \end{bmatrix}.
\]

Solution. The vectors \( u_1 \) and \( u_2 \) are orthogonal to each other. First we compute \( \hat{y} \), the orthogonal projection of \( y \) onto \( W \):

\[
\hat{y} = \frac{1 + 9 - 10}{1 + 9 + 4} u_1 + \frac{5 + 3 + 20}{25 + 1 + 16} u_2 = 0 u_1 + \frac{2}{3} u_2 = \begin{bmatrix} 10/3 \\ 2/3 \\ 8/3 \end{bmatrix}.
\]

Let \( v = y - \hat{y} = (-7/3, 7/3, 7/3) \). Then \( y = \hat{y} + v \) is the required sum.

Exercise 8: Let \( W = \text{Span}\{u_2, u_2\} \). Write \( y \) as the sum of a vector in \( W \) and a vector orthogonal to \( W \). Here

\[
y = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}, \quad u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad u_2 = \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}.
\]

Solution. The vectors \( u_1 \) and \( u_2 \) are orthogonal to each other. First we compute \( \hat{y} \), the orthogonal projection of \( y \) onto \( W \):

\[
\hat{y} = \frac{-1 + 4 + 3}{1 + 1 + 1} u_1 + \frac{1 + 12 - 6}{1 + 9 + 4} u_2 = 2 u_1 + \frac{1}{2} u_2 = \begin{bmatrix} 3/2 \\ 7/2 \\ 1 \end{bmatrix}.
\]

Let \( v = y - \hat{y} = (-5/2, 1/2, 2) \). Then \( y = \hat{y} + v \) is the required sum.

Section 6.4

Exercise 10: Find an orthogonal basis for the column space of

\[
A = [x_1, x_2, x_3] = \begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & -3 \end{bmatrix}.
\]
**Solution.** We apply the Gram-Schmidt process. We let \( v_1 = x_1 \). Also,
\[
v_2 = x_2 - \frac{-6 - 24 - 2 - 4}{1 + 9 + 1 + 1}\ v_1 = x_2 + 3v_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix}.
\]

Finally, we should let
\[
v_3 = x_3 - \frac{-6 + 9 + 6 - 3}{1 + 9 + 1 + 1}\ v_1 - \frac{18 + 3 + 6 + 3}{9 + 1 + 1 + 1}\ v_2 = x_3 - \frac{1}{2} v_1 - \frac{5}{2} v_2 = \begin{bmatrix} -1 \\ -1 \\ 3 \\ -1 \end{bmatrix}.
\]

A routine checking shows that the obtained vectors \( v_1, v_2, v_3 \) are indeed orthogonal.

---

**Section 6.5**

**Exercise 12:** Find (a) the orthogonal projection of \( b \) into \( \text{Col} \ A \) and (b) a least-square solution of \( Ax = b \). Here
\[
A = [v_1 \ v_2 \ v_3] = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 5 \\ 6 \\ 6 \end{bmatrix}.
\]

**Solution.** It is easy to check that the columns of \( A \) are orthogonal to each other. (In particular, they are linearly independent.) Hence, we can use the standard formulas for finding the orthogonal projection of \( b \) onto \( \text{Col} \ A \):
\[
b = \frac{2 + 5 - 6}{1 + 1 + 1} v_1 + \frac{2 + 6 + 6}{1 + 1 + 1} v_2 + \frac{-5 + 6 - 6}{1 + 1 + 1} v_1 = \frac{1}{3} (v_1 + 14v_2 - 5v_3) = \begin{bmatrix} 5 \\ 2 \\ 3 \\ 6 \end{bmatrix}.
\]

This answers (a). Since the columns of \( A \) are linearly independent, the least-square solution \( x \) is unique and we already know the weights, namely \( x = (1/3, 14/3, -5/3) \).

As an check, one can compute \( b - \hat{b} = (-3, 3, 3, 0) \) and see that it is indeed orthogonal to each \( v_i \).

---

**Section 6.6**

**Exercise 4:** Find the equation \( y = \beta_0 + \beta_1 x \) of the least-squares line that fits best the given data points:
\[
(2, 3), \ (3, 2), \ (5, 1), \ (6, 0).
\]
Solution. We construct the design matrix and the observation vector:

\[
X = \begin{bmatrix}
1 & 2 \\
1 & 3 \\
1 & 5 \\
1 & 6
\end{bmatrix}, \quad y = \begin{bmatrix}
3 \\
2 \\
1 \\
0
\end{bmatrix}.
\]

We want to find the least-squares solution to \(X\bar{\beta} = y\). The normal equation is

\[
X^T X \bar{\beta} = X^T y.
\]

We have

\[
X^T X = \begin{bmatrix} 4 & 16 \\ 16 & 74 \end{bmatrix}, \quad X^T y = \begin{bmatrix} 6 \\ 17 \end{bmatrix}.
\]

It is probably easier to compute first the inverse

\[
(X^T X)^{-1} = \frac{1}{4 \cdot 74 - 16^2} \begin{bmatrix} 74 & -16 \\ -16 & 4 \end{bmatrix} = \frac{1}{40} \begin{bmatrix} 74 & -16 \\ -16 & 4 \end{bmatrix} = \frac{1}{20} \begin{bmatrix} 37 & -8 \\ -8 & 2 \end{bmatrix}.
\]

Hence, the least-squares solution is

\[
\bar{\beta} = \frac{1}{20} \begin{bmatrix} 37 & -8 \\ -8 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 17 \end{bmatrix} = \begin{bmatrix} 43/10, -7/10 \end{bmatrix}.
\]

Thus the least-squares line is \(y = 4.3 - 0.7x\).

Exercise 6: Let \(X\) be the design matrix corresponding to a least-squares fit of a parabola to data \((x_1, y_1), \ldots, (x_n, y_n)\). Suppose that \(x_1, x_2, x_3\) are distinct. Explain why there is only one parabola that fits the data best, in a least-squares sense.

Solution. It is enough to prove that the columns of \(X\) are linearly independent, since then \(X^T X\) is invertible and the unique least-squares solution is \((X^T X)^{-1} X^T y\).

Let us remove Row 1 from any other row of \(X\):

\[
X = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ \vdots \end{bmatrix} \sim \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 - x_1 & x_2^2 - x_1^2 \\ 1 & x_3 - x_1 & x_3^2 - x_1^2 \\ \vdots \end{bmatrix}
\]

Calculations show that the determinant

\[
\begin{vmatrix} x_2 - x_1 & x_2^2 - x_1^2 \\ x_3 - x_1 & x_3^2 - x_1^2 \end{vmatrix} = (x_2 - x_1)(x_3 - x_1) \begin{vmatrix} 1 & x_2 + x_1 \\ 1 & x_3 + x_1 \end{vmatrix} = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2).
\]

This is non-zero since \(x_1, x_2, x_3\) are distinct by the assumption. Thus this \(2 \times 2\)-matrix is invertible and has 2 pivot columns. This means that if we continue the row reduction of \(X\), then we get 3 pivots. Thus the columns of \(X\) are independent, as required.
Section 7.1

Exercise 14: Orthogonally diagonalize matrix

\[ A = \begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix} \]

Solution. The characteristic equation is \((1 - \lambda)^2 - 25 = 0\). The roots are \(-4\) and \(6\). We have

\[ A + 4I = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad A - 6I = \begin{bmatrix} -5 & 5 \\ -5 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}. \]

The corresponding eigenvectors are \(v_1 = (-1, 1)\) and \(v_2 = (1, 1)\). They are orthogonal as we expected them to be. Let us normalize them, by multiplying each by \(1/\sqrt{2}\). We let

\[ P = \frac{1}{\sqrt{2}} [v_1 \ v_2] = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}, \quad D = \begin{bmatrix} -4 & 0 \\ 0 & 6 \end{bmatrix}. \]

Since \(P\) is orthogonal, we have \(P^{-1} = P^T\).

\[ A = PDP^{-1} = PD^TP \]

is the required orthogonal diagonalization.

Exercise 22: Orthogonally diagonalize matrix \(A\), given that its eigenvalues are 0 and 2, where

\[ A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}. \]

Solution. Let us find the eigenvectors corresponding to the eigenvalue \(\lambda = 0\), which amounts to finding a basis for the \(\text{Nul} \ A\). We have

\[ A \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \]

We have one free variable \(x_4\), so \(\text{Nul} \ A\) is 1-dimensional and it spanned by \(v_1 = (0, -1, 0, 1)\). Let us immediately normalize \(v_1\) by replacing it with \(v_1 = (0, -1/\sqrt{2}, 0, 1/\sqrt{2})\).

For \(\lambda = 2\), we obtain

\[ A - 2I = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \]
Here $x_1, x_3, x_4$ are free; the general solution to $(A - 2I)x = 0$ is

$$x = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$ 

Luckily for us, the obtained 3 vectors are already orthogonal, so we just normalize them, having

$$v_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}.$$ 

Now we let

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1/\sqrt{2} & 0 & 0 & 1/\sqrt{2} \\ 0 & 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 0 & 1/\sqrt{2} \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix}.$$ 

Since $P$ is orthogonal, we have $P^{-1} = P^T$. 

$$A = PDP^{-1} = PDP^T$$ 

is the required orthogonal diagonalization. 

**Exercise 32:** Suppose that $A = PRP^{-1}$, where $P$ is orthogonal and $R$ is upper triangular. Show that if $A$ is symmetric, then $R$ is symmetric and hence is actually a diagonal matrix. 

**Solution.** By the assumptions we have $P^{-1} = P^T$ and $A^T = A$. This means that

$$PRP^T = A = A^T = (PRP^T)^T = (P^T)^TR^TP^T = PRT^TP.$$ 

But $P^TP = P^{-1}$ are inverses of each other. So if we multiply the obtained identity by $P^{-1}$ on left and by $P$ on right, we obtain $R = R^T$. Thus $R$ is symmetric. Since all entries of $T$ below the main diagonal are zeros, by symmetry all entries above the main diagonals are zeros too. So $R$ is also diagonal. 

**Section 7.2**

**Exercise 10:** Let $Q(x_1, x_2) = 9x_1^2 - 8x_1x_2 + 3x_2^2$. Classify the type of $Q$ and make a change of variables $x = Py$ that eliminates all cross-product terms. 

**Solution.** The matrix of $Q$ is

$$A = \begin{bmatrix} 9 & -4 \\ -4 & 3 \end{bmatrix}.$$
First, we find the eigenvalues of $A$. The characteristic polynomial is

$$P_A(\lambda) = (9 - \lambda)(3 - \lambda) - 16.$$  

Its roots are $\lambda_1 = 1$ and $\lambda_2 = 11$. Both are positive so $Q$ is **positive definite**. (Of course, it is also positive semidefinite but of all types of $Q$ we usually mention the one which is most precise.)

Let us compute the corresponding unit eigenvectors. We have

$$A - I = \begin{bmatrix} 8 & -4 \\ -4 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1/2 \\ 0 & 0 \end{bmatrix}.$$  

We can take a vector $(1, 2)$. After normalizing it by $1/\sqrt{5}$, $v_1 = (1/\sqrt{5}, 2/\sqrt{5})$. Next,

$$A - 11I = \begin{bmatrix} -2 & -4 \\ -4 & -8 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}.$$  

Here we take $v_2 = (2/\sqrt{5}, -1/\sqrt{5})$. The vectors $v_1$ and $v_2$ are orthogonal as they should be (and each of norm 1). Hence, we take

$$P = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 11 \end{bmatrix}.$$  

Then the transformation $x = Py$ transforms $Q$ into $y_1^2 + 11y_2^2$.

Now, it could be a good idea to check this by hand. We have $x_1 = y_1/\sqrt{5} + 2y_2/\sqrt{5}$ and $x_2 = 2y_1/\sqrt{5} - y_2/\sqrt{5}$. Then

$$Q = \frac{9}{5}(y_1^2 + 2y_2/\sqrt{5})^2 - 8(y_1/\sqrt{5} + 2y_2/\sqrt{5})(2y_1/\sqrt{5} - y_2/\sqrt{5}) + 3(2y_1/\sqrt{5} - y_2/\sqrt{5})^2$$

$$= \frac{1}{5}(9y_1^2 + 4y_1y_2 + 4y_2^2) - 8(2y_1^2 + 3y_1y_2 - 2y_2^2) + 3(4y_1^2 - 4y_1y_2 + y_2^2) = y_1^2 + 11y_2^2.$$  

So, everything is OK!

**Section 7.3**

**Exercise 6**: Let $Q(x) = 7x_1^2 + 3x_2^2 + 3x_1x_2$. Find a) the maximum of $Q(x)$ subject to the constraint $x^T x = 1$, b) a unit vector where this maximum is attained, and c) the maximum of $Q(x)$ subject to the constraints $x^T x = 1$ and $x^T u = 0$.

**Solution.** The matrix of $Q$ is

$$A = \begin{bmatrix} 7 & 3/2 \\ 3/2 & 3 \end{bmatrix}.$$  

Its eigenvalues are $\lambda_1 = 5/2$ and $\lambda_2 = 15/2$ with eigenvectors $v_2 = (-1, 3)$ and $v_1 = (3, 1)$. Hence the answer to a) is $15/2$; the answer to b) is $v_1/\|v_1\| = (3/\sqrt{10}, 1/\sqrt{10})$; the answer to c) is $5/2$.  

Section 7.4

Exercise 10: Find an SVD of

\[
A = \begin{bmatrix}
4 & -2 \\
2 & -1 \\
0 & 0 \\
\end{bmatrix}.
\]

Solution. In Step 1 we orthogonally diagonalize

\[
A^T A = \begin{bmatrix}
20 & -10 \\
-10 & 5 \\
\end{bmatrix}.
\]

Its eigenvalues are \( \lambda_1 = 25 \) and \( \lambda_2 = 0 \). (We list them in decreasing order.) The corresponding normalized eigenvectors are \( v_1 = (-2/\sqrt{5}, 1/\sqrt{5}) \) and \( v_2 = (1/\sqrt{5}, 2/\sqrt{5}) \). By the way, although we are not required to do this, we can immediately write an orthogonal diagonalization

\[
A^T A = V D V^T,
\]

where

\[
V = [v_1 \ v_2] = \begin{bmatrix}
-2/\sqrt{5} & 1/\sqrt{5} \\
1/\sqrt{5} & 2/\sqrt{5}
\end{bmatrix}, \quad D = \begin{bmatrix}
25 & 0 \\
0 & 0
\end{bmatrix}.
\]

At this point it is a good idea to check if \( A^T A V = V D \).

The singular values of \( A \) are \( 5, 0, 0 \). In Step 2 we take the same matrix \( V \) as above. The middle matrix \( \Sigma \) should have dimensions \( 3 \times 2 \), the same as those of \( A \), so we just add a row of zeros:

\[
\Sigma = \begin{bmatrix}
5 & 0 \\
0 & 0 \\
0 & 0
\end{bmatrix}.
\]

In Step 3 we construct \( U \). We see that \( A \) (or equivalently \( \Sigma \)) has rank \( r = 1 \). So the first column of \( U \) is

\[
\mathbf{u}_1 = \frac{A v_1}{\| A v_1 \|} = \frac{(-10/\sqrt{5}, -5/\sqrt{5}, 0)}{\sqrt{20 + 5}} = \begin{bmatrix}
-2/\sqrt{5} \\
-1/\sqrt{5} \\
0
\end{bmatrix}.
\]

We choose the remaining columns \( \mathbf{u}_2 \) and \( \mathbf{u}_3 \) of \( U \) so that \( U \) is orthogonal. This is the same as finding an orthonormal basis of

\[
\text{Nul}(\mathbf{u}_1^T) = \{ \mathbf{u} \in \mathbb{R}^3 \mid \mathbf{u}_1^T \mathbf{u} = 0 \}.
\]

We have

\[
\mathbf{u}_1^T \sim \begin{bmatrix} 1 & 1/2 & 0 \end{bmatrix},
\]

so a basis for \( \text{Nul}(\mathbf{u}_1^T) \) is \( \mathbf{w}_2 = (1, -2, 0) \) and \( \mathbf{w}_3 = (0, 0, 1) \). Luckily for us these vectors are already orthogonal. (If they were not, then we would have to apply the Gram-Schmidt process to them.) So, it remains only to normalize them, obtaining \( \mathbf{u}_2 = (1/\sqrt{5}, -2/\sqrt{5}, 0) \) and \( \mathbf{u}_3 = (0, 0, 1) \). We take

\[
U = [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3].
\]

Thus the required SVD of \( A \) is

\[
A = \begin{bmatrix}
-2/\sqrt{5} & 1/\sqrt{5} & 0 \\
-1/\sqrt{5} & -2/\sqrt{5} & 0 \\
0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
5 & 0 \\
0 & 0 \\
0 & 0
\end{bmatrix} \begin{bmatrix}
-2/\sqrt{5} & 1/\sqrt{5} \\
1/\sqrt{5} & 2/\sqrt{5}
\end{bmatrix}.
\]
Finally, we are done!