Problem 1: Do problem 4 in section 6.7 (pg. 360) in the book.

Solution (10 points)

a) We have

\[ A^T A = A A^T = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \]

This matrix has eigenvalues satisfying \( \lambda^2 - 3\lambda + 1 = 0 \), so it has eigenvalues \( \lambda_1 = \frac{3}{2} + \frac{\sqrt{5}}{2} \) and \( \lambda_2 = \frac{3}{2} - \frac{\sqrt{5}}{2} \). Its eigenvectors form the nullspace of

\[ A^T A - \frac{3 + \sqrt{5}}{2} I = \begin{bmatrix} (1 - \sqrt{5})/2 & 1 \\ 1 & -(1 + \sqrt{5})/2 \end{bmatrix} \]

This has nullspace generated by \((2, \sqrt{5} - 1)\). Since the eigenvectors of \( A^T A \) must be perpendicular, we know that another eigenvector is \((\sqrt{5} - 1, -2)\) (which we could also find directly). The normalized eigenvector matrix is

\[ S = \frac{1}{\sqrt{10 - 2\sqrt{5}}} \begin{bmatrix} 2 & \sqrt{5} - 1 \\ \sqrt{5} - 1 & -2 \end{bmatrix} \]

b) We construct the singular value decomposition \( A = U \Sigma V^H \). First, we choose the matrix \( V \) to be the eigenvector matrix for \( A^T A \); that is, it is just the \( S \) we found in part a. The matrix \( \Sigma \) is the 2x2 matrix with the square roots of the eigenvalues of \( A^T A \) on the diagonal:

\[ \Sigma = \begin{bmatrix} \sqrt{\frac{3 + \sqrt{5}}{2}} & 0 \\ 0 & \sqrt{\frac{3 - \sqrt{5}}{2}} \end{bmatrix} \]

Finally, we find \( U \) via the equation \( AV = U \Sigma \). We can’t skip directly to \( U = S \). It is true that \( U \) will be an eigenvector matrix for \( A A^T \), but we must pick the eigenvectors correctly! In this case the only choice in unit eigenvectors of \( A A^T \) is the sign. Even so, we must have the relationship \( A = U \Sigma V^H \), and if we get the sign of the vectors of \( U \) backwards this will not be true.
Let \( v_i \) and \( u_i \) be the ith columns of \( V \) and \( U \). We know \( u_1 \) is either \( v_1 \) or \(-v_1 \), and similarly for \( u_2 \). The question is just which way around it is. We start with \( v_1 \):

\[
Av_1 = \sqrt{\frac{3 + \sqrt{5}}{2}} u_1
\]

\[
\frac{1}{\sqrt{10 - 2\sqrt{5}}} \begin{bmatrix} \sqrt{5} + 1 \\ 2 \end{bmatrix} = \sqrt{\frac{3 + \sqrt{5}}{2}} u_1
\]

so

\[
u_1 = \sqrt{\frac{1}{10 + 2\sqrt{5}}} \begin{bmatrix} \sqrt{5} + 1 \\ 2 \end{bmatrix} = \sqrt{\frac{1}{10 - 2\sqrt{5}}} \begin{bmatrix} 2 \\ \sqrt{5} - 1 \end{bmatrix}
\]

This is the same vector as \( v_1 \). Here \( Av_1 = \sigma_1 u_1 \) is an eigenvector equation for \( A \), since \( \sigma_1 \) is an eigenvalue of \( A \). So \( v_1 \) keeps the same sign.

For \( v_2 \) we find:

\[
Av_2 = \sqrt{\frac{3 - \sqrt{5}}{2}} u_2
\]

\[
\frac{1}{\sqrt{10 - 2\sqrt{5}}} \begin{bmatrix} \sqrt{5} - 3 \\ \sqrt{5} - 1 \end{bmatrix} = \sqrt{\frac{3 - \sqrt{5}}{2}} u_2
\]

We already know that \( u_2 \) is either \( v_2 \) or \(-v_2 \). However \( v_2 \) has negative second component, and \( u_2 \) has negative first component, meaning that the sign has switched. Here \( Av_2 = \sigma_2 u_2 \) is not an eigenvector equation, since \( \sigma_2 = -\lambda_2 \). So we need to switch the sign of \( u_2 \) as well.

In the end, we get the SVD:

\[
U = \frac{1}{\sqrt{10 - 2\sqrt{5}}} \begin{bmatrix} 2 \\ -(\sqrt{5} - 1) \end{bmatrix}
\]

\[
\Sigma = \begin{bmatrix} \sqrt{\frac{3 + \sqrt{5}}{2}} \\ 0 \end{bmatrix}
\]

\[
V = \frac{1}{\sqrt{10 - 2\sqrt{5}}} \begin{bmatrix} 2 \\ \sqrt{5} - 1 \end{bmatrix}
\]

It is almost the diagonalization of \( A \), but not quite. Since one of the eigenvalues of \( A \) is negative, it can’t appear in \( \Sigma \). We must switch its sign, and we compensate by switching the sign of the eigenvector in \( U \). As you might guess from this problem,
the SVD for a positive definite matrix is its diagonalization – see the last problem of this pset.

**Problem 2:** Do problem 7 in section 6.7 (pg. 360).

**Solution** (10 points)
Here

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

Here the eigenvalue equation is \((1 - \lambda)(\lambda^2 - 3\lambda + 1) - (1 - \lambda) = 0\). Factoring out the \((1 - \lambda)\), we get \((1 - \lambda)\lambda(\lambda - 3) = 0\), so the eigenvalues are 3, 1, 0. Remember, when we do the SVD we always put 0 eigenvalues last! This is important.

The first eigenvector is the nullspace of 

\[ A^TA - 3I = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & -2 \end{bmatrix} \]

By inspection we see that this has basis \((1, 2, 1)\). Similarly, the second eigenvector is the nullspace of

\[ A^TA - I = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \]

By inspection this has basis \((1, 0, -1)\). Finally, the last eigenvector is the nullspace of \(A^TA\), and by inspection we see this is \((1, -1, 1)\). Putting this all together, we get

a normalized eigenvector matrix

\[ S = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix} \]

Now we repeat this for

\[ AA^T = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \]

This has eigenvalues given by \(\lambda^2 - 4\lambda + 3 = 0\), so the eigenvalues are 3 and 1. An eigenvector for 3 is \((1, 1)/\sqrt{2}\), and for 1 is \((1, -1)/\sqrt{2}\).
Finally, we find the SVD. As before, we set $V = S$ that we found above. We find the $2 \times 3$ matrix $\Sigma$ by taking the square roots of the eigenvalues (either for $A^T A$ or $AA^T$, both will work):

$$\Sigma = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Finally, we find $U$ using the equations $Av_i = \sigma_i u_i$. As before, we know that $U$ is an eigenvector matrix for $AA^T$, but we must choose the correct one. Here the unit eigenvectors are determined up to sign.

Calculating:

$$Av_1 = \begin{bmatrix} \frac{3}{\sqrt{6}} \\ \frac{3}{\sqrt{6}} \end{bmatrix} = \sqrt{3}u_1$$

So we set $u_1 = (1,1)/\sqrt{2}$. Similarly

$$Av_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = u_2$$

So we get the SVD:

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$V = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix}$$

Finally we check by multiplying it all out:

$$U\Sigma V^H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}$$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 2/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix}$$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 2/\sqrt{2} & 2/\sqrt{2} & 0 \\ 0 & 2/\sqrt{2} & 2/\sqrt{2} \end{bmatrix}$$

$$= A$$
Note that the last row of $V^H$ didn’t affect anything. This is typical when we get eigenvalues of 0; they shouldn’t factor in to the multiplication at all.

**Problem 3:** Do problem 9 in section 6.7 (pg. 361).

**Solution** (5 points)

First note that $A$ must have dimensions 3 by 4. If $A$ has rank one, so does $A^T A$. This means that only one eigenvalue of $A^T A$ is not 0, so $\Sigma$ has the form

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Because we only have one non-zero entry in $\Sigma$, we also only get one non-trivial equation $Av_1 = \sigma_1 u_1$. Of course this must be the equation given in the problem $Av = 12u$. So, the first column of $U$ is $u$, and the first column of $V$ is $v$.

When we multiply out $A = U \Sigma V^T$, most of it will cancel because of the 0 entries in $\Sigma$. In fact, the only non-zero part will come from the first columns of $U$ and $V$ (see part a of the next problem). So $A = 12uv^T$. You don’t need to multiply it out, but if you do you get

$$A = \begin{bmatrix} 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \\ 2 & 2 & 2 & 2 \end{bmatrix}$$

The only singular value is given by the equation, namely, $\sigma_1 = 12$.

We could also have done this problem by noting that any rank 1 matrix has the form $xy^T$ for some vectors $x$ and $y$, and using the equation to calculate $x$ and $y$ explicitly.

**Problem 4:**

a) Do problem 11 in section 6.7 (pg. 361).

b) Do problem 16 in section 6.7 (pg. 361).

**Solution** (5+5 points)

a) In brief, the SVD expresses $A$ as a sum of $r$ rank one matrices because of the block form of multiplication (see page 60). The block form of multiplication is a general fact, so the only thing to write down is why $\Sigma$ has the effect that it does.

So, note that if there are more columns than rows, then multiplication by $\Sigma$ rescales the rows of the matrix $V$ and cuts off the bottom ones. Similarly, if there are more rows than columns, multiplication by $\Sigma$ rescales the columns of $U$ and cuts
off the last ones. Either way, using the block picture of matrix multiplication, we find \( U \Sigma V^T \) as a sum of rank one matrices

\[
U \Sigma V^T = u_1 \sigma_1 v_1^T + \ldots + u_r \sigma_r v_r^T
\]

b) One might hope that if \( A \) were a square matrix, the SVD for \( A + I \) would involve \( \Sigma + I \) in analogy to the diagonalization equation. However, if we were to use \( \Sigma + I \) in the SVD, we would get \( U(\Sigma + I)V^H = A + UV^H \neq A + I \). The problem is that \( \Sigma \) is the square root of the eigenvalues of \( A^TA \). Substituting \( A + I \) in gives \( (A^T + I)(A + I) = A^T A + A^T + A + I \), and the eigenvalues don’t work out right in general.

Problem 5: Do problem 6 in section 7.1 (pg. 368).

Solution (10 points)

a) This \( T \) does not satisfy either criterion. For example, if \( v = (1, 0, 0) \) and \( w = (0, 1, 0) \), then \( T(v + w) = (1, 1, 0)/\sqrt{2} \neq (1, 0, 0) + (0, 1, 0) \) and \( T(2v) = (1, 0, 0) \neq 2(1, 0, 0) \).

b) This satisfies both; it is a linear transformation. In fact, it is the linear transformation from \( \mathbb{R}^3 \) to \( \mathbb{R} \) given by multiplying by the matrix \([1, 1, 1]\).

c) This again satisfies both; it is the linear transformation from \( \mathbb{R}^3 \) to \( \mathbb{R}^3 \) given by the matrix

\[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 2 & 0 \\
0 & 0 & 3
\end{bmatrix}
\]

d) This satisfies neither criterion. For example, if \( v = (-1, 0, 0) \) and \( w = (2, 0, 0) \), then \( T(v + w) = 1 \neq 0 + 2 \) and \( T(-v) = 1 \neq -1(0) \).

Problem 6: Do problem 12 in section 7.1 (pg. 369).

Solution (10 points)

The quickest way to do each of these is to write the given vector as a linear combination of the basis \((1, 1)\) and \((2, 0)\). To find the coefficients in the new basis, we multiply by the change-of-base matrix

\[
\begin{bmatrix}
1 & 2 \\
1 & 0
\end{bmatrix}^{-1} = \begin{bmatrix}
0 & 1 \\
1/2 & -1/2
\end{bmatrix}
\]
a) Because
\[
\begin{bmatrix}
0 & 1 \\
1/2 & -1/2
\end{bmatrix}
\begin{bmatrix}
2 \\
2
\end{bmatrix} =
\begin{bmatrix}
2 \\
0
\end{bmatrix}
\]
we see that \((2, 2) = 2(1, 1) + 0(2, 0)\). (Of course we could have seen this more easily directly.) So \(T((2, 2)) = 2T(1, 1) + 0T(2, 0) = 2(2, 2) = (4, 4)\).

b) Because
\[
\begin{bmatrix}
0 & 1 \\
1/2 & -1/2
\end{bmatrix}
\begin{bmatrix}
3 \\
1
\end{bmatrix} =
\begin{bmatrix}
1 \\
1
\end{bmatrix}
\]
we see that \((3, 1) = (1, 1) + (2, 0)\). So \(T((3, 1)) = T(1, 1) + T(2, 0) = (2, 2) + (0, 0) = (2, 2)\).

c) Because
\[
\begin{bmatrix}
0 & 1 \\
1/2 & -1/2
\end{bmatrix}
\begin{bmatrix}
-1 \\
1
\end{bmatrix} =
\begin{bmatrix}
1 \\
-1
\end{bmatrix}
\]
we see that \((-1, 1) = (1, 1) - (2, 0)\). So \(T((-1, 1)) = T(1, 1) - T(2, 0) = (2, 2)\).

d) Because
\[
\begin{bmatrix}
0 & 1 \\
1/2 & -1/2
\end{bmatrix}
\begin{bmatrix}
a \\
b
\end{bmatrix} =
\begin{bmatrix}
b \\
a/2 - b/2
\end{bmatrix}
\]
we see that \((a, b) = b(1, 1) + \frac{a-b}{2}(2, 0)\). So \(T((a, b)) = bT(1, 1) + \frac{a-b}{2}T(2, 0) = b(2, 2)\).

Problem 7: Do problems 5 and 7 in section 7.2 (pg. 380-381).

Solution (5+5 points)

Problem 5: \(T\) is a linear transformation from the three-dimensional space \(V\) to the three-dimensional space \(W\). Once we choose a basis for \(V\) and \(W\) we can associate a (unique) matrix to \(T\). Remember, we form the the \(i\)th column of \(A\) by putting in \(T(v_i)\) in terms of \(w_i\). For example, because \(T(v_1) = w_2\), the first column must be \(\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}^T\). Thus \(T\) must have the matrix
\[
A = \begin{bmatrix}
0 & 1 & 1 \\
1 & 0 & 0 \\
0 & 1 & 1
\end{bmatrix}
\]

Problem 7: Since \(T(v_2) = T(v_3)\) (and there are no other linear relations), the nullspace of \(T\) has basis \(v_2 - v_3\). That is, \(T(c(v_2 - v_3)) = c(T(v_2) - T(v_3)) = 0\). This corresponds to the column vector \(\begin{bmatrix} 0 & 1 & -1 \end{bmatrix}^T\), which one can check for \(A\) easily.
The complete solution to $T(v) = w_2$ is the particular solution plus the nullspace. Since a particular solution is $v_1$, the complete solution is all vectors of the form $v_1 + c(v_2 - v_3)$, or in vectors $[1, 0, 0]^T + c[0, 1, -1]^T$.

**Problem 8:** Do problem 16 in section 7.2 (pg. 381).

**Solution** (10 points)

a) This is just the matrix
\[
\begin{bmatrix}
  r & s \\
  t & u
\end{bmatrix}
\]

Remember that the first column of a matrix is where $(1, 0)$ goes, and the second column is where $(0, 1)$ goes.

b) This is the change-of-base matrix that is the inverse of the change we just did:
\[
\begin{bmatrix}
  a & b \\
  c & d
\end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix}
  d & -b \\
  -c & a
\end{bmatrix}
\]

You can check by hand!

c) Of course we can’t do this when $ad - bc = 0$, that is, we can’t do this if the vectors are dependent. If they are in the same direction, we must also get vectors in the same direction after doing $T$.

**Problem 9:** Do problem 28 in section 7.2 (pg. 382).

**Solution** (5 points)

Repeating the statement: suppose we have an invertible linear transformation. Then pick any basis $v_1, \ldots, v_n$ of $V$, and pick the basis $w_i = T(v_i)$ of $W$. Then of course with these bases $T$ corresponds to the identity matrix.

The question is why we need $T$ to be invertible for this to work. If $T$ is not invertible, then in fact the $T(v_i)$ can’t form a basis because they will be linearly dependent. This is because if $T$ is not invertible, then there is a vector $a_1v_1 + \ldots + a_nv_n$ in the nullspace (and not all of the $a_i$ are 0). That is,

$$T(a_1v_1 + \ldots + a_nv_n) = a_1T(v_1) + \ldots + a_nT(v_n) = 0$$

This gives a linear dependence relation between the $T(v_i)$. 8
If $T$ is invertible, then the $T(v_i)$ must be linearly independent, for precisely the same reason; if there were a linear relation, then $T$ would have to have a non-trivial nullspace.

**Problem 10:** Do problem 13 in section 7.4 (pg. 398).

**Solution**  (10 points)

Here $A$ is a 1 by 3 matrix, so $U$ will be 1 by 1 and $V$ will be 3 by 3. We start by finding $V$ and $\Sigma$. Note that $A^TA$ will have eigenvector $[3, 4, 0]^T$ with eigenvalue 25, and then two perpendicular eigenvectors each with eigenvalue 0. We can find these eigenvectors by taking the nullspace of $A$: it has special solutions $[-4/3, 1, 0]^T$ and $[0, 0, 1]^T$. Remember that we must renormalize these vectors when forming $V$. So we have

$$V = \begin{bmatrix} 3/5 & -4/5 & 0 \\ 4/5 & 3/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The singular value $\sigma_1 = 5$ is the square root of the eigenvalue. Finally, since $U$ is a unit 1 by 1 vector, it must be either $[1]$ or $[-1]$, and using $Av_1 = \sigma_1u_1$ shows that it is $[1]$. Writing it all down, we get

$$A = [1] \begin{bmatrix} 5 & 0 & 0 \\ 3/5 & -4/5 & 0 \\ 4/5 & 3/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}^H$$

The pseudoinverse $A^+ = V\Sigma^+U^H$. Writing it down, we get

$$A^+ = \begin{bmatrix} 3/5 & -4/5 & 0 \\ 4/5 & 3/5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/5 \\ 0 \\ 0 \end{bmatrix}^H$$

The product $AA^+$ is projection onto the column space of $A$. However, the column space of $A$ is just $[c]$. So we should expect to get the identity 1 by 1 matrix:

$$AA^+ = U\Sigma V^H V \Sigma^+ U^H = U\Sigma \Sigma^+ U^H = UU^H = [1]$$

The other way round, $A^+A$ is projection onto the row space of $A$. Calculating, we get

$$A^+A = V\Sigma^+ U^H U\Sigma V^H = V \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} V^H$$

$$= v_1v_1^T$$
and since $v_1$ is a unit vector, this is just projection onto the space generated by $v_1$, namely, the row space of $A$.

**Problem 11:** Do problem 16 in section 7.4 (pg. 399).

**Solution** (10 points)

The SVD will equal the diagonalization $QΛQ^T$ when $A$ is symmetric positive semi-definite. (The answer “positive definite” is acceptable, since that is what the phrasing would lead you to believe.)

Let’s prove it by diagonalizing $A^TA$ to find $V$ and $Σ$. Suppose that $A$ is symmetric positive semidefinite - then it has non-negative real eigenvalues and orthonormal eigenvectors. Write the diagonalization $A = QΛQ^T$. We have $A^TA = A^2$, so the diagonalization is $A^TA = QΛ^2Q^T$. Thus $V = Q$. Also, because all of the eigenvalues are non-negative, taking the square roots of the entries of $Λ^2$ returns $Λ$. So $Σ = Λ$. Finally, $U = AVΣ^{-1} = Q$ as well.

Note: if $A$ weren’t positive semidefinite, then the square roots of the diagonal of $Λ^2$ wouldn’t give us $Λ$ because some of the signs would be switched. $U$ would then be $Q$ but with some of the signs of the vectors switched to compensate.