Math 2270 § 1.	**Final Exam**	Name: Practice Problems
Treibergs		November 30, 2015

Half of the final exam will be comprehensive. The other half will focus on material since the last midterm exam. These problems are representative of this last part of the course.

1. *Show that the orthogonal projection of a vector* $\mathbf{y} \in \mathbf{R}^n$ *onto a line* L *through the origin does not depend on the choice of the nonzero vector* $\mathbf{u}$ *in* L *used in the formula for* $\widehat{\mathbf{y}}$. *(Problem 346[31] of the text.)*

We shall show that replacing $\mathbf{u}$ by another vector $\mathbf{w} = c\mathbf{u}$ where c is an unspecified nonzero scalar results in the same vector $\widehat{\mathbf{y}}$. The projection is given by the formula

$$\widehat{\mathbf{y}} = \operatorname{proj}_L \mathbf{y} = \frac{\mathbf{y} \bullet \mathbf{u}}{\mathbf{u} \bullet \mathbf{u}}\mathbf{u}.$$

Let us see what we get if we replace $\mathbf{u}$ by $\mathbf{w}$.

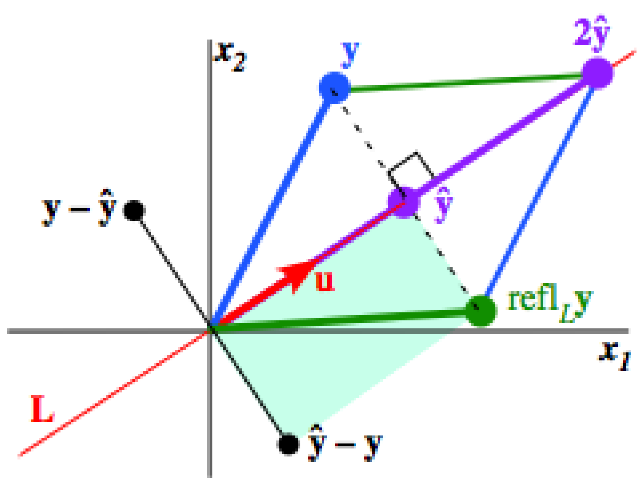
$$\frac{\mathbf{y} \bullet \mathbf{w}}{\mathbf{w} \bullet \mathbf{w}}\mathbf{w} = \frac{\mathbf{y} \bullet c\mathbf{u}}{c\mathbf{u} \bullet c\mathbf{u}}c\mathbf{u} = \frac{c^2(\mathbf{y} \bullet \mathbf{u})}{c^2(\mathbf{u} \bullet \mathbf{u})}\mathbf{u} = \frac{\mathbf{y} \bullet \mathbf{u}}{\mathbf{u} \bullet \mathbf{u}}\mathbf{u} = \widehat{\mathbf{y}}$$

which is the same for any c.

2. *Given* $\mathbf{u} \neq \mathbf{0}$ *in* $\mathbf{R}^n$, *let* $L = \operatorname{span}\{\mathbf{u}\}$. *For* $\mathbf{y} \in \mathbf{R}^n$, *the reflection of* $\mathbf{y}$ *in* L *is the point* $\operatorname{refl}_L \mathbf{y}$ *as in the figure defined by*

$$\operatorname{refl}_L \mathbf{y} = 2 \operatorname{proj}_L \mathbf{y} - \mathbf{y}$$

Show that the mapping $F : \mathbf{y} \mapsto \operatorname{refl}_L \mathbf{y}$ *is a linear transformation. (Problem 346[34] of the text.)*



Let's find a formula for $\operatorname{refl}_L \mathbf{y}$. Using the formula for $\operatorname{proj}_L \mathbf{y}$ we see that

$$F(\mathbf{y}) = \operatorname{refl}_L \mathbf{y} = 2 \operatorname{proj}_L \mathbf{y} - \mathbf{y} = 2\frac{\mathbf{u} \bullet \mathbf{y}}{\mathbf{u} \bullet \mathbf{u}}\mathbf{u} - \mathbf{y}$$

Now, a linear transformation must preserve vector addition and scalar multiplication. Choosing $\mathbf{y}, \mathbf{z} \in \mathbf{R}^n$ and $c \in \mathbf{R}$ we see that

$$F(\mathbf{y} + \mathbf{z}) = 2\frac{\mathbf{u} \bullet (\mathbf{y} + \mathbf{z})}{\mathbf{u} \bullet \mathbf{u}}\mathbf{u} - (\mathbf{y} + \mathbf{z}) = \left(2\frac{\mathbf{u} \bullet \mathbf{y}}{\mathbf{u} \bullet \mathbf{u}}\mathbf{u} - \mathbf{y}\right) + \left(2\frac{\mathbf{u} \bullet \mathbf{z}}{\mathbf{u} \bullet \mathbf{u}}\mathbf{u} - \mathbf{z}\right) = F(\mathbf{y}) + F(\mathbf{z});$$

$$F(c\mathbf{y}) = 2\frac{\mathbf{u} \bullet (c\mathbf{y})}{\mathbf{u} \bullet \mathbf{u}}\mathbf{u} - (c\mathbf{y}) = c\left(2\frac{\mathbf{u} \bullet \mathbf{y}}{\mathbf{u} \bullet \mathbf{u}}\mathbf{u} - \mathbf{y}\right) = cF(\mathbf{y}).$$

3. *Let A be the 8×4 matrix given. Find the closest point to $\mathbf{y} = (1,1,1,1,1,1,1,1)^T$ in* $\operatorname{Col} A$*. How far from* $\operatorname{Col} A$ *is it? (Text problems 346[36] and 354[25]).*

$$A = \begin{pmatrix} -6 & -3 & 6 & 1 \\ -1 & 2 & 1 & -6 \\ 3 & 6 & 3 & -2 \\ 6 & -3 & 6 & -1 \\ 2 & -1 & 2 & 3 \\ -3 & 6 & 3 & 2 \\ -2 & -1 & 2 & -3 \\ 1 & 2 & 1 & 6 \end{pmatrix}$$

Let U be the matrix whose columns have been normalized to length one. Let $\mathbf{a}_j$ for $j = 1,2,3,4$ be the columns of A. To compute inner products of the columns we compute $A^T A$ whose entiries are $\mathbf{a}_i \bullet \mathbf{a}_j$

$$A^T A = \begin{pmatrix} 100 & 0 & 0 & 0 \\ 0 & 100 & 0 & 0 \\ 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 100 \end{pmatrix}$$

Hence A has orthogonal columns. Also each column has length $\|\mathbf{a}_i\| = \sqrt{\mathbf{a}_i \bullet \mathbf{a}_i} = 10$. Let $U = \frac{1}{10} A$ be the matrix with lengths of the columns normalized to length one

$$U = \begin{pmatrix} -.6 & -.3 & .6 & .1 \\ -.1 & .2 & .1 & .6 \\ .3 & .6 & .3 & -.2 \\ .6 & -.3 & .6 & -.1 \\ .2 & -.1 & .2 & .3 \\ -.3 & .6 & .3 & .2 \\ -.2 & -.1 & .2 & -.3 \\ .1 & .2 & .1 & .6 \end{pmatrix}$$

Note that $\operatorname{Col} A = \operatorname{Col} U$. The projection of $\mathbf{w} \in \mathbf{R}^n$ to $\operatorname{Col} A$ is given by

$$\operatorname{proj}_{\operatorname{Col} A} \mathbf{w} = (\mathbf{u}_1 \bullet \mathbf{w})\,\mathbf{u}_1 + \cdots + (\mathbf{u}_4 \bullet \mathbf{w})\,\mathbf{u}_4 = U \begin{pmatrix} \mathbf{u}_1 \bullet \mathbf{w} \\ \mathbf{u}_2 \bullet \mathbf{w} \\ \mathbf{u}_3 \bullet \mathbf{w} \\ \mathbf{u}_4 \bullet \mathbf{w} \end{pmatrix} = UU^T\mathbf{w}.$$

Applying this to the given $\mathbf{y}$ we find

$$\widehat{\mathbf{y}} = UU^T\mathbf{y} = \begin{pmatrix} 1.2 \\ 0.4 \\ 1.2 \\ 1.2 \\ 0.4 \\ 1.2 \\ 0.4 \\ 0.4 \end{pmatrix}$$

$\widehat{\mathbf{y}} = \operatorname{proj}_{\operatorname{Col} A} \mathbf{y}$ is the closest point in $\operatorname{Col} A$ to $\mathbf{y}$. The orthogonal component is

$$\mathbf{z} = \mathbf{y} - \widehat{\mathbf{y}} = \begin{pmatrix} -.2 \\ .6 \\ -.2 \\ -.2 \\ .6 \\ -.2 \\ .6 \\ .6 \end{pmatrix}$$

Thus the distance of $\mathbf{y}$ to $\operatorname{Col} A$ is

$$\operatorname{dist}(\mathbf{y}, \operatorname{Col} A) = \|\mathbf{z}\| = \sqrt{\mathbf{z} \bullet \mathbf{z}} = \sqrt{1.6} = 1.264911.$$

4. *Find an orthonormal basis for* Col A *using the Gram-Schmidt process. Find the QR factorization of A. (Text problem 360[12,16].)*

$$A = \begin{pmatrix} 1 & 3 & 5 \\ -1 & -3 & 1 \\ 0 & 2 & 3 \\ 1 & 5 & 2 \\ 1 & 5 & 8 \end{pmatrix}$$

Denote the columns $A = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3]$. Then the Gram-Schmidt algorithm to find an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\mathbf{v}_3\}$ is

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{a}_1 \\ \mathbf{v}_2 &= \mathbf{a}_2 - \frac{\mathbf{a}_2 \bullet \mathbf{v}_1}{\mathbf{v}_1 \bullet \mathbf{v}_1}\mathbf{v}_1 \\ \mathbf{v}_3 &= \mathbf{a}_3 - \frac{\mathbf{a}_3 \bullet \mathbf{v}_1}{\mathbf{v}_1 \bullet \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{a}_3 \bullet \mathbf{v}_2}{\mathbf{v}_2 \bullet \mathbf{v}_2}\mathbf{v}_2 \end{aligned}$$

In terms of the given vectors

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix}; \qquad \mathbf{v}_2 = \begin{pmatrix} 3 \\ -3 \\ 2 \\ 5 \\ 5 \end{pmatrix} - \frac{16}{4}\begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}$$

$$\mathbf{v}_3 = \begin{pmatrix} 5 \\ 1 \\ 3 \\ 2 \\ 8 \end{pmatrix} - \frac{14}{4}\begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \frac{12}{8}\begin{pmatrix} -1 \\ 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 0 \\ -3 \\ 3 \end{pmatrix}$$

This says $\mathbf{a}_1 = \mathbf{v}_1$, $\mathbf{a}_2 = 4\mathbf{v}_1 + \mathbf{v}_2$ and $\mathbf{a}_3 = \frac{7}{2}\mathbf{v}_1 + \frac{3}{2}\mathbf{v}_2 + \mathbf{v}_3$. These equations written in matrix form and then pulling out lengths yields

$$A = \begin{pmatrix} 1 & 3 & 5 \\ -1 & -3 & 1 \\ 0 & 2 & 3 \\ 1 & 5 & 2 \\ 1 & 5 & 8 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 3 \\ -1 & 1 & 3 \\ 0 & 2 & 0 \\ 1 & 1 & -3 \\ 1 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 4 & \frac{7}{2} \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2} & -\frac{1}{2\sqrt{2}} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2\sqrt{2}} & \frac{1}{2} \\ 0 & \frac{2}{2\sqrt{2}} & 0 \\ \frac{1}{2} & \frac{1}{2\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2\sqrt{2}} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2\sqrt{2} & 0 \\ 0 & 0 & 6 \end{pmatrix} \begin{pmatrix} 1 & 4 & \frac{7}{2} \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2} & -\frac{1}{2\sqrt{2}} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2\sqrt{2}} & \frac{1}{2} \\ 0 & \frac{2}{2\sqrt{2}} & 0 \\ \frac{1}{2} & \frac{1}{2\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2\sqrt{2}} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 & 8 & 7 \\ 0 & 2\sqrt{2} & 3\sqrt{2} \\ 0 & 0 & 6 \end{pmatrix} = QR.$$

The book recommends another procedure. First determine Q and then find R from Q and A. The lengths are $\|\mathbf{v}_1\|^2 = 4$, $\|\mathbf{v}_2\|^2 = 8$, and $\|\mathbf{v}_1\|^2 = 36$. Dividing each column of the matrix $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$ by its length gives the matrix Q

$$Q = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2\sqrt{2}} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2\sqrt{2}} & \frac{1}{2} \\ 0 & \frac{2}{2\sqrt{2}} & 0 \\ \frac{1}{2} & \frac{1}{2\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2\sqrt{2}} & \frac{1}{2} \end{pmatrix}$$

Since $Q^TQ = I$, the matrix $R = Q^T(QR) = Q^TA$ is obtained by

$$R = Q^TA = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{2}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & 0 & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 3 & 5 \\ -1 & -3 & 1 \\ 0 & 2 & 3 \\ 1 & 5 & 2 \\ 1 & 5 & 8 \end{pmatrix} = \begin{pmatrix} 2 & 8 & 7 \\ 0 & 2\sqrt{2} & 3\sqrt{2} \\ 0 & 0 & 6 \end{pmatrix}$$

5. *Find the least-squares solution of $A\mathbf{x} = \mathbf{b}$ in three ways. What is the least-squares error of approximation?*

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 1 \\ 3 \\ 8 \\ 2 \end{pmatrix}$$

The first method is to solve the normal equation for $\widehat{\mathbf{x}}$.

$$A^TA\widehat{\mathbf{x}} = A^T\mathbf{b}.$$

Because columns of A are independent, A^TA is invertible

$$A^TA = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 1 \\ 2 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

$$A^T\mathbf{b} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 8 \\ 2 \end{pmatrix} = \begin{pmatrix} 14 \\ 4 \\ 8 \end{pmatrix}$$

Solve by row reducing the augmented matrix $[A^TA, A^T\mathbf{b}]$

$$\begin{pmatrix} 4 & 2 & 1 & 14 \\ 2 & 2 & 0 & 4 \\ 1 & 0 & 1 & 8 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 1 & 8 \\ 2 & 2 & 0 & 4 \\ 4 & 2 & 1 & 14 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 1 & 8 \\ 0 & 2 & -2 & -12 \\ 0 & 2 & -3 & -18 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 1 & 8 \\ 0 & 1 & -1 & -6 \\ 0 & 0 & 1 & 6 \end{pmatrix}$$

Thus $\hat{x}_3 = 6$, $\hat{x}_2 = -6 + \hat{x}_3 = 0$ and $\hat{x}_1 = 8 - \hat{x}_3 = 2$. Thus the projection of $\mathbf{b}$ onto $\operatorname{Col} A$ is

$$\widehat{\mathbf{y}} = A\widehat{\mathbf{x}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 6 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 8 \\ 2 \end{pmatrix}$$

The perpendicular vector is

$$\mathbf{b} - A\widehat{\mathbf{x}} = \begin{pmatrix} 1 \\ 3 \\ 8 \\ 2 \end{pmatrix} - \begin{pmatrix} 2 \\ 2 \\ 8 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

so the least-squares error is $\|\mathbf{b} - A\widehat{\mathbf{x}}\| = \sqrt{2}$.

The second method is to replace the columns of A by orthonormal column vectors of $\tilde{A}$. Then compute $\widehat{\mathbf{b}}$ and $\widetilde{\mathbf{x}}$ using the orthonormal vectors. Using Gram-Schmidt process,

$$\mathbf{v}_1 = \mathbf{a}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\mathbf{v}_2 = \mathbf{a}_2 - \frac{\mathbf{a}_2 \bullet \mathbf{v}_1}{\mathbf{v}_1 \bullet \mathbf{v}_1} \mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \frac{2}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$$

$$\mathbf{v}_3 = \mathbf{a}_3 - \frac{\mathbf{a}_3 \bullet \mathbf{v}_1}{\mathbf{v}_1 \bullet \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{a}_3 \bullet \mathbf{v}_2}{\mathbf{v}_2 \bullet \mathbf{v}_2} \mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{\left(-\frac{1}{2}\right)}{1} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$$

Then we normalixe the lengths to get the matrix Q with orthonormal columns. The matrix was constructed so that $\operatorname{Col} A = \operatorname{Col} Q$.

$$Q = \left[\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}, \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} \right] = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{\sqrt{2}} \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 1 \\ 3 \\ 8 \\ 2 \end{pmatrix}$$

Now the projection is given by

$$\widehat{\mathbf{b}} = \operatorname{proj}_{\operatorname{Col} Q} \mathbf{b} = (\mathbf{q}_1 \bullet \mathbf{b})\,\mathbf{q}_1 + (\mathbf{q}_2 \bullet \mathbf{b})\,\mathbf{q}_2 + (\mathbf{q}_3 \bullet \mathbf{b})\,\mathbf{q}_3 = Q\widetilde{\mathbf{x}}.$$

Compting,

$$\widetilde{\mathbf{x}} = \begin{pmatrix} \mathbf{q}_1 \bullet \mathbf{b} \\ \mathbf{q}_2 \bullet \mathbf{b} \\ \mathbf{q}_3 \bullet \mathbf{b} \end{pmatrix} = Q^T\mathbf{b} = \begin{pmatrix} 7 \\ -3 \\ 3\sqrt{2} \end{pmatrix}, \qquad \widehat{\mathbf{b}} = Q\widetilde{\mathbf{x}} = QQ^T\mathbf{b} = \begin{pmatrix} 2 \\ 2 \\ 8 \\ 2 \end{pmatrix}$$

which is the same $\widehat{\mathbf{b}}$ as before.

The third method is to find $\widehat{\mathbf{x}}$ using the QR decomposition. We have the matrix Q with orthonormal columns such that $\operatorname{span}\{\mathbf{a}_1\} = \operatorname{span}\{\mathbf{q}_1\}$, $\operatorname{span}\{\mathbf{a}_1, \mathbf{a}_2\} = \operatorname{span}\{\mathbf{q}_1, \mathbf{q}_2\}$ and $\operatorname{span}\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\} = \operatorname{span}\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$. Since $Q^TA = Q^TQR = IR = R$ we have

$$R = Q^TA = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

Because $A\widehat{\mathbf{x}} = QR\widehat{\mathbf{x}} = \widehat{\mathbf{b}} = QQ^T\mathbf{b}$ it means that we solve the triangular system $R\widehat{\mathbf{x}} = Q^T\mathbf{b}$ by back substitution.

$$\begin{pmatrix} 2 & 1 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 8 \\ 2 \end{pmatrix} = \begin{pmatrix} 7 \\ -3 \\ 3\sqrt{2} \end{pmatrix}$$

Solving yields $\hat{x}_3 = 6$, $\hat{x}_2 = -3 + \frac{1}{2}\hat{x}_3 = 0$ and $\hat{x}_1 = \frac{1}{2}(7 - \hat{x}_2 - \frac{1}{2}\hat{x}_3) = 2$, which is the same $\widehat{\mathbf{x}}$ as before.

6. *Orthogonally diagonalize the matrix A. (Text problem 401[19].)*

$$A = \begin{pmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{pmatrix}$$

The characteristic polynomial is

$$\begin{vmatrix} 3-\lambda & -2 & 4 \\ -2 & 6-\lambda & 2 \\ 4 & 2 & 3-\lambda \end{vmatrix} \begin{aligned} &= (3-\lambda)^2(6-\lambda) - 16 - 16 - 4(3-\lambda) - 4(3-\lambda) - 16(6-\lambda) \\ &= (9 - 6\lambda + \lambda^2)(6-\lambda) - 152 + 24\lambda \\ &= 54 - 45\lambda + 12\lambda^2 - \lambda^3 - 152 + 24\lambda \\ &= -(98 + 21\lambda - 12\lambda^2 + \lambda^3) \\ &= -(\lambda-7)^2(\lambda+2). \end{aligned}$$

For the eigenvalue $\lambda_1 = -2$, we row reduce $A - \lambda_1 I$

$$\begin{pmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{pmatrix} \to \begin{pmatrix} 1 & -.4 & .8 \\ 0 & 7.2 & 3.6 \\ 0 & 3.6 & 1.8 \end{pmatrix} \to \begin{pmatrix} 1 & -.4 & .8 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathbf{v}_1 = \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix}$$

x_3 is free so take an eigenvector to be $x_3 = 2$, $x_2 = -1$ and $x_1 = .4x_2 - .8x_3 = -2$. For the eigenvalue $\lambda_2 = 7$, we find two independent eigenvectors by inspection

$$(A - \lambda_2 I)[\mathbf{v}_2, \mathbf{v}_3] = \begin{pmatrix} -4 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -4 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 2 & 0 \\ 0 & 1 \end{pmatrix}$$

The $\lambda_2 = 7$ eigenvectors $\mathbf{v}_2$ and $\mathbf{v}_3$ are not orthogonal. Apply Gram-Schmidt process to get

$$\mathbf{w}_2 = \mathbf{v}_2 = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \qquad \mathbf{w}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3 \bullet \mathbf{w}_2}{\mathbf{w}_3 \bullet \mathbf{w}_2}\mathbf{w}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{(-1)}{5} \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{4}{5} \\ \frac{2}{5} \\ 1 \end{pmatrix}$$

We normalize by dividing columns by their lengths. $\|\mathbf{v}_1\| = 3$, $\|\mathbf{w}_2\| = \sqrt{5}$ and $\|\mathbf{W}_3\| = \frac{3\sqrt{5}}{5}$. We get an orthogonal matrix of eigenvectors

$$P = \begin{pmatrix} -\frac{2}{3} & -\frac{\sqrt{5}}{5} & \frac{4\sqrt{5}}{15} \\ -\frac{1}{3} & \frac{2\sqrt{5}}{5} & \frac{2\sqrt{5}}{15} \\ \frac{2}{3} & 0 & \frac{\sqrt{5}}{3} \end{pmatrix}$$

so $P^T P = I$. We check that $P^T AP = D$. Indeed it checks

$$AP = \begin{pmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{pmatrix} \begin{pmatrix} -\frac{2}{3} & -\frac{\sqrt{5}}{5} & \frac{4\sqrt{5}}{15} \\ -\frac{1}{3} & \frac{2\sqrt{5}}{5} & \frac{2\sqrt{5}}{15} \\ \frac{2}{3} & 0 & \frac{\sqrt{5}}{3} \end{pmatrix} = \begin{pmatrix} \frac{4}{3} & -\frac{7\sqrt{5}}{5} & \frac{28\sqrt{5}}{15} \\ \frac{2}{3} & \frac{14\sqrt{5}}{5} & \frac{14\sqrt{5}}{15} \\ -\frac{4}{3} & 0 & \frac{7\sqrt{5}}{3} \end{pmatrix}$$

$$PD = \begin{pmatrix} -\frac{2}{3} & -\frac{\sqrt{5}}{5} & \frac{4\sqrt{5}}{15} \\ -\frac{1}{3} & \frac{2\sqrt{5}}{5} & \frac{2\sqrt{5}}{15} \\ \frac{2}{3} & 0 & \frac{\sqrt{5}}{3} \end{pmatrix} \begin{pmatrix} -2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{pmatrix} = \begin{pmatrix} \frac{4}{3} & -\frac{7\sqrt{5}}{5} & \frac{28\sqrt{5}}{15} \\ \frac{2}{3} & \frac{14\sqrt{5}}{5} & \frac{14\sqrt{5}}{15} \\ -\frac{4}{3} & 0 & \frac{7\sqrt{5}}{3} \end{pmatrix}$$

7. *Suppose that both A and B are orthogonally diagonalizable and $AB = BA$. Explain why AB is also orthogonally diagonalizable. (Text problem 401[30].)*

   This follows from the deep Spectral Theorem.

   Theorem 1. *Let A be an $n \times n$ matrix. Then A is orthogonally diagonalizable if and only if it is symmetric.*

   The fact that an orthogonally diagonalizable A is symmetric is not stated as part of the book's version, but is easy to prove. A orthogonally diagonalizable means there is an orthogonal matrix P such that $P^T AP = D$. This implies $A = PDP^T$. But then

$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$$

   where we have used transpose of transpose is the identity $(P^T)^T = P$ and transpose of diagonal is itself $D^T = D$.

   To prove the claim it suffices to show that AB is a symmetric matrix. Indeed

$$(AB)^T = B^T A^T = BA = AB$$

   where we used the fact that since both are orthogonally diagonalizable we have A and B are symmetric, and then the hypothesis $AB = BA$. By the Spectral Theorem, AB is now orthogonally diagonalizable.

8. *Classify the quadratic form as positive definite, positive semidefinite, etc. Make a change of variables* $\mathbf{x} = P\mathbf{y}$ *that transforms the quadratic form into one with no cross product terms. Write the new quadratic form. (Text problem 408[16].)*

$$Q(\mathbf{x}) = 4x_1^2 + 4x_2^2 + 4x_3^2 + 4x_4^2 + 8x_1x_2 + 8x_3x_4 - 6x_1x_4 + 6x_2x_3$$

The form may be written $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ where

$$A = \begin{pmatrix} 4 & 4 & 0 & -3 \\ 4 & 4 & 3 & 0 \\ 0 & 3 & 4 & 4 \\ -3 & 0 & 4 & 4 \end{pmatrix}$$

Let us orthogonally diagonalize A to find P. The eigenvalues are found by expanding the first row

$$\begin{vmatrix} 4-\lambda & 4 & 0 & -3 \\ 4 & 4-\lambda & 3 & 0 \\ 0 & 3 & 4-\lambda & 4 \\ -3 & 0 & 4 & 4-\lambda \end{vmatrix} =$$

$$= (4-\lambda)\begin{vmatrix} 4-\lambda & 3 & 0 \\ 3 & 4-\lambda & 4 \\ 0 & 4 & 4-\lambda \end{vmatrix} - 4\begin{vmatrix} 4 & 3 & 0 \\ 0 & 4-\lambda & 4 \\ -3 & 4 & 4-\lambda \end{vmatrix} + 3\begin{vmatrix} 4 & 4-\lambda & 3 \\ 0 & 3 & 4-\lambda \\ -3 & 0 & 4 \end{vmatrix}$$

$$= (4-\lambda)^4 - 25(4-\lambda)^2 - 4[4(4-\lambda)^2 - 36 - 64] + 3[48 - 3(4-\lambda)^2 + 27]$$
$$= (4-\lambda)^4 - 25(4-\lambda)^2 - 16(4-\lambda)^2 + 625 - 9(4-\lambda)^2$$
$$= (4-\lambda)^4 - 50(4-\lambda)^2 + 625 = [(4-\lambda)^2 - 25]^2 = (9-\lambda)^2(\lambda+1)^2$$

For $\lambda_1 = 9$, we get the eigenvectors by row reducing the matrix $A - \lambda_1 I$

$$\begin{pmatrix} -5 & 4 & 0 & -3 \\ 4 & -5 & 3 & 0 \\ 0 & 3 & -5 & 4 \\ -3 & 0 & 4 & -5 \end{pmatrix} \to \begin{pmatrix} 1 & -.8 & 0 & .6 \\ 0 & -1.8 & 3 & -2.4 \\ 0 & 3 & -5 & 4 \\ 0 & -2.4 & 4 & -3.2 \end{pmatrix} \to \begin{pmatrix} 1 & -.8 & 0 & .6 \\ 0 & 3 & -5 & 4 \\ 0 & 3 & -5 & 4 \\ 0 & 3 & -5 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & -.8 & 0 & .6 \\ 0 & 3 & -5 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The eigenvectors are thus

$$\mathbf{v}_1 = \begin{pmatrix} 4 \\ 5 \\ 3 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -5 \\ -4 \\ 0 \\ 3 \end{pmatrix}, \quad \mathbf{w}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \bullet \mathbf{w}_1}{\mathbf{w}_1 \bullet \mathbf{w}_1}\mathbf{w}_1 = \begin{pmatrix} -5 \\ -4 \\ 0 \\ 3 \end{pmatrix} - \frac{(-40)}{50}\begin{pmatrix} 4 \\ 5 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} -1.8 \\ 0 \\ 2.4 \\ 3 \end{pmatrix}$$

They are not orthogonal. Applying Gram-Schmid we have $\mathbf{w}_1 = \mathbf{v}_1$ and $\mathbf{w}_2$ above. For $\lambda_2 = -1$, we get the eigenvectors by row reducing the matrix $A - \lambda_2 I$

$$\begin{pmatrix} 5 & 4 & 0 & -3 \\ 4 & 5 & 3 & 0 \\ 0 & 3 & 5 & 4 \\ -3 & 0 & 4 & 5 \end{pmatrix} \to \begin{pmatrix} 1 & .8 & 0 & -.6 \\ 0 & 1.8 & 3 & 2.4 \\ 0 & 3 & 5 & 4 \\ 0 & 2.4 & 4 & 3.2 \end{pmatrix} \to \begin{pmatrix} 1 & .8 & 0 & -.6 \\ 0 & 3 & 5 & 4 \\ 0 & 3 & 5 & 4 \\ 0 & 3 & 5 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & .8 & 0 & -.6 \\ 0 & 3 & 5 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The eigenvectors are thus

$$\mathbf{v}_3 = \begin{pmatrix} 4 \\ -5 \\ 3 \\ 0 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 5 \\ -4 \\ 0 \\ 3 \end{pmatrix}, \quad \mathbf{w}_4 = \mathbf{v}_4 - \frac{\mathbf{v}_4 \bullet \mathbf{w}_3}{\mathbf{w}_3 \bullet \mathbf{w}_3}\mathbf{w}_3 = \begin{pmatrix} 5 \\ -4 \\ 0 \\ 3 \end{pmatrix} - \frac{40}{50}\begin{pmatrix} 4 \\ -5 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 1.8 \\ 0 \\ -2.4 \\ 3 \end{pmatrix}$$

They are not orthogonal. Applying Gram-Schmid we have $\mathbf{w}_3 = \mathbf{v}_3$ and $\mathbf{w}_4$ above. Dividing by the norms of the columns of $[\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4]$ vectors, we finally obtain the orthogonal matrix

$$P = \frac{\sqrt{2}}{10}\begin{pmatrix} 4 & -3 & 4 & 3 \\ 5 & 0 & -5 & 0 \\ 3 & 4 & 3 & -4 \\ 0 & 5 & 0 & 5 \end{pmatrix}$$

Changing variables according to $\mathbf{x} = P\mathbf{y}$, the quadratic form in the $\mathbf{y}$ variable is

$$Q(\mathbf{y}) = (P\mathbf{y})^T AP\mathbf{y} = \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T B\mathbf{y}$$

where

$$B = P^T A P = \begin{pmatrix} 9 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

which means that in these coordinates, the quadratic form is

$$Q(\mathbf{y}) = 9y_1^2 + 9y_2^2 - y_3^2 - y_4^2$$

This means that the form is classified as $\boxed{\text{indefinite}}$.

9. *Let* $A = \left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right)$ *be such that* $\det A \neq 0$ *and let* $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$. *Show that if* λ_1 *and* λ_2 *are eigenvalues of* A*, then the characteristic polynomial of* A *can be written in two ways*

$$\chi(\lambda) = \det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)$$

Use this fact to show that $\operatorname{trace}(A) = a + c = \lambda_1 + \lambda_2$ *and* $\det A = \lambda_1\lambda_2$. *Verify the following statements: (Problems 408[23,24] of the text.)*

(a) *Q is positive definite if* $\det A > 0$ *and* $a > 0$.
(b) *Q is negative definite if* $\det A > 0$ *and* $a < 0$.
(c) *Q is indefinite if* $\det A < 0$.

Writing out the determinant we find

$$\det(A - \lambda I) = \begin{vmatrix} a - \lambda & b \\ b & c - \lambda \end{vmatrix} = (a - \lambda)(c - \lambda) - b^2 = \lambda^2 - (a + c)\lambda + (ac - b^2)$$

The roots of this polynomial are λ_1 and λ_2. By the fundamental theorem of algebra, every polynomial may be factored into linear factors thus there is a constant k such that

$$\chi(\lambda) = k(\lambda_1 - \lambda)(\lambda_2 - \lambda) = k[\lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2]$$

The powers of λ in both expressions have to agree, which implies

$$k = 1; \qquad a + c = \lambda_1 + \lambda_2; \qquad \det A = ac - b^2 = \lambda_1\lambda_2.$$

Using this, we can give the classification of Q without knownig the eigenvalues. If $\det A < 0$ then we know right away that $\lambda_1\lambda_2 < 0$ by the equation so one eigenvalue is positive and the other is negative. In principal axis coordinates, Q has directions in which it is positive and directions in which it's negative, therefore Q is indefinite.

If $\det A > 0$ then λ_1 and λ_2 are both positive or both negative. Which of these cases holds can be determined by knowing the sign of $\operatorname{trace} A$ which has the same sign as both eigenvalues. It can also be determined from knowing just the sign of a. $\det A > 0$ implies

$$ac = b^2 + \det A > 0$$

If $a > 0$ then we must have $c > 0$ also which implies $\operatorname{trace} A = a + c > 0$ so both eigenvalues are positive and Q is positive definite. If $a < 0$ then this says $c < 0$ and $\operatorname{trace} A < 0$ so Q is negative definite.

10. *Find the maximum value of* $Q(\mathbf{x})$ *subject to the constraint* $\mathbf{x}^T\mathbf{x} = 1$. *Fnd a unit vector* $\mathbf{u}$ *where this maximum is attained. Find the maximum of* $Q(\mathbf{x})$ *subject to the constraints* $\mathbf{x}^T\mathbf{x} = 1$ *and* $\mathbf{x}^T\mathbf{u} = 0$. *(Text problem 415[4].)*

$$Q(\mathbf{x}) = 3x_1^2 + 3x_2^2 + 5x_3^2 + 6x_1x_2 + 2x_1x_3 + 2x_2x_3$$

The quadratic form may be written $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ where

$$A = \begin{pmatrix} 3 & 3 & 1 \\ 3 & 3 & 1 \\ 1 & 1 & 5 \end{pmatrix}$$

The maximum corresponds to the maximum eigenvalue. Computing the characteristic equation we find

$$A = \begin{vmatrix} 3-\lambda & 3 & 1 \\ 3 & 3-\lambda & 1 \\ 1 & 1 & 5-\lambda \end{vmatrix} \begin{aligned} &= (3-\lambda)^2(5-\lambda) + 6 - 2(3-\lambda) - 9(5-\lambda) \\ &= (9 - 6\lambda + \lambda^2)(5-\lambda) + 11\lambda - 45 \\ &= 45 - 30\lambda + 5\lambda^2 - 9\lambda + 6\lambda^2 - \lambda^3 + 11\lambda - 45 \\ &= -28\lambda + 11\lambda^2 - \lambda^3 = -\lambda(\lambda^2 - 11\lambda + 28) \\ &= -\lambda(\lambda-4)(\lambda-7) \end{aligned}$$

We find the $\lambda_1 = 7$ eigenvector by row reducing $A - \lambda_1 I$

$$\begin{pmatrix} -4 & 3 & 1 \\ 3 & -4 & 1 \\ 1 & 1 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & -2 \\ 3 & -4 & 1 \\ -4 & 3 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 \\ 0 & -7 & 7 \\ 0 & 7 & -7 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Thus $\mathbf{v}_1 = (1,1,1)^T$. We find the $\lambda_2 = 4$ eigenvector by row reducing $A - \lambda_2 I$

$$\begin{pmatrix} -1 & 3 & 1 \\ 3 & -1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} -1 & 3 & 1 \\ 0 & 8 & 4 \\ 0 & 4 & 2 \end{pmatrix} \to \begin{pmatrix} -1 & 3 & 1 \\ 0 & 8 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

Thus $\mathbf{v}_2 = (-1,-1,2)^T$.

The maximum of $Q(\mathbf{x})$ is 7 corresponding to the largest eigenvalue $\lambda_1 = 7$. The unit vector for which it occurs is a normalized eigenvector $\mathbf{u} = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right)^T$.

The maximum of $Q(\mathbf{x})$ subject to the constraints $\mathbf{x}^T\mathbf{x} = 1$ and $\mathbf{x}^T\mathbf{u} = 0$ is 4 and corresponds to the second largest eigenvalue $\lambda_2 = 4$ and is taken at the unit vector $\mathbf{w} = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \left(-\frac{\sqrt{6}}{6}, -\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right)^T$ which satisfies both constraints.

11. *Let A be an $n \times n$ symmetric matrix. Let M and m denote the maximum and minimum values the quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ where $\mathbf{x}^T\mathbf{x} = 1$, and denote the corresponding eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_n$. Show that for any number t between m and M, there is a unit vector $\mathbf{x}$ such that $Q(\mathbf{x}) = t$. (Text problem 415[13].)*

   (The problem actually tells you how to do this.) The number t may be obtained as a convex combination
$$t = (1-\alpha)M + \alpha m$$
   for some number $\alpha \in [0,1]$. Consider the vector
$$\mathbf{x} = \sqrt{1-\alpha}\,\mathbf{u}_1 + \sqrt{\alpha}\,\mathbf{u}_n.$$
   It satisfies the constraint.
$$\begin{aligned}\mathbf{x}^T\mathbf{x} &= (\sqrt{1-\alpha}\,\mathbf{u}_1 + \sqrt{\alpha}\,\mathbf{u}_n)^T(\sqrt{1-\alpha}\,\mathbf{u}_1 + \sqrt{\alpha}\,\mathbf{u}_n)\\ &= (1-\alpha)\mathbf{u}_1^T\mathbf{u}_1 + 2\sqrt{\alpha(1-\alpha)}\,\mathbf{u}_1^T\mathbf{u}_n + \alpha\,\mathbf{u}_n^T\mathbf{u}_n\\ &= (1-\alpha) + 0 + \alpha = 1.\end{aligned}$$
   Because $A\mathbf{u}_1 = M\mathbf{u}_1$ and $A\mathbf{u}_n = m\mathbf{u}_n$ the quadratic form takes the correct value.
$$\begin{aligned}Q(\mathbf{x}) &= \mathbf{x}^T A\mathbf{x}\\ &= (\sqrt{1-\alpha}\,\mathbf{u}_1 + \sqrt{\alpha}\,\mathbf{u}_n)^T A(\sqrt{1-\alpha}\,\mathbf{u}_1 + \sqrt{\alpha}\,\mathbf{u}_n)\\ &= (1-\alpha)\mathbf{u}_1^T A\mathbf{u}_1 + \sqrt{\alpha(1-\alpha)}\,(\mathbf{u}_1^T A\mathbf{u}_n + \mathbf{u}_n^T A\mathbf{u}_1) + \alpha\,\mathbf{u}_n^T A\mathbf{u}_n\\ &= (1-\alpha)M\,\mathbf{u}_1^T\mathbf{u}_1 + \sqrt{\alpha(1-\alpha)}\,(m\mathbf{u}_1^T\mathbf{u}_n + M\mathbf{u}_n^T\mathbf{u}_1) + \alpha m\,\mathbf{u}_n^T\mathbf{u}_n\\ &= (1-\alpha)M + 0 + \alpha m = t.\end{aligned}$$

12. *Find the singular value decomposition of the matrix A. (Text problem 425[11].)*
$$A = \begin{pmatrix} -3 & 1\\ 6 & -2\\ 6 & -2 \end{pmatrix}$$
   We find
$$A^T A = \begin{pmatrix} -3 & 6 & 6\\ 1 & -2 & -2 \end{pmatrix}\begin{pmatrix} -3 & 1\\ 6 & -2\\ 6 & -2 \end{pmatrix} = \begin{pmatrix} 81 & -27\\ -27 & 9 \end{pmatrix}$$
   The characteristic polynomial is
$$\det(A^T A - \lambda I) = (81-\lambda)(9-\lambda) - 729 = \lambda^2 - 90\lambda$$
   so that the eigenvalues are $\lambda_1 = 90$ and $\lambda_2 = 0$ so the singular values are $\sigma_1 = \sqrt{\lambda_1} = 3\sqrt{10}$ and $\sigma_2 = \sqrt{\lambda_2} = 0$. The eigenvectors are found by inspection
$$(A^T A - \lambda_1 I)\widetilde{\mathbf{v}_1} = \begin{pmatrix} -9 & -27\\ -27 & -81 \end{pmatrix}\begin{pmatrix} -3\\ 1 \end{pmatrix}, \qquad (A^T A - \lambda_2 I)\widetilde{\mathbf{v}_2} = \begin{pmatrix} 81 & -27\\ -27 & 9 \end{pmatrix}\begin{pmatrix} 1\\ 3 \end{pmatrix}.$$

Normalizing, we find the orthogonal matrix

$$V = [\mathbf{v}_1, \mathbf{v}_2] = \left[\frac{\widetilde{\mathbf{v}_1}}{\|\widetilde{\mathbf{v}_1}\|}, \frac{\widetilde{\mathbf{v}_2}}{\|\widetilde{\mathbf{v}_2}\|}\right] = \frac{1}{\sqrt{10}}\begin{pmatrix} -3 & 1 \\ 1 & 3 \end{pmatrix}$$

The singular values give the "diagonal" matrix

$$\Sigma = \begin{pmatrix} 3\sqrt{10} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

The corresponding vector

$$\mathbf{u}_1 = \frac{A\mathbf{u}_1}{\|A\mathbf{u}_1\|} = \begin{pmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{pmatrix}\begin{pmatrix} -\frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix} = \frac{1}{3\sqrt{10}}\begin{pmatrix} \sqrt{10} \\ -2\sqrt{10} \\ -2\sqrt{10} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ -\frac{2}{3} \end{pmatrix}$$

generates $\operatorname{Col} A$ since A has rank one. $\mathbf{u}_2$ and $\mathbf{u}_3$ complete $\mathbf{u}_1$ to an orthonormal basis of $\mathbf{R}^3$. Two vectors orthogonal to $\mathbf{u}_1$ satisfy the system $x_1 - 2x_2 - 2x_3 = 0$. Two solutions are

$$\mathbf{w}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \qquad \mathbf{w}_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \qquad \widetilde{\mathbf{w}_2} = \mathbf{w}_2 - \frac{\mathbf{w}_1 \bullet \mathbf{w}_2}{\mathbf{w}_1 \bullet \mathbf{w}_1}\mathbf{w}_1 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} - \frac{4}{5}\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{2}{5} \\ -\frac{4}{5} \\ 1 \end{pmatrix}$$

Applying Gram-Schmidt, we can make the second orthogonal to the first. By normalizing we obtain

$$U = \left[\frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \frac{\widetilde{\mathbf{w}_2}}{\|\widetilde{\mathbf{w}_2}\|}\right] = \begin{pmatrix} \frac{1}{3} & \frac{2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} \\ -\frac{2}{3} & \frac{1}{\sqrt{5}} & -\frac{4}{3\sqrt{5}} \\ -\frac{2}{3} & 0 & \frac{\sqrt{5}}{3} \end{pmatrix}$$

We have obtained the SVD for $A = U\Sigma V^T$. We check

$$U\Sigma = \begin{pmatrix} \frac{1}{3} & \frac{2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} \\ -\frac{2}{3} & \frac{1}{\sqrt{5}} & -\frac{4}{3\sqrt{5}} \\ -\frac{2}{3} & 0 & \frac{\sqrt{5}}{3} \end{pmatrix}\begin{pmatrix} 3\sqrt{10} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \sqrt{10} & 0 \\ -2\sqrt{10} & 0 \\ -2\sqrt{10} & 0 \end{pmatrix},$$

$$AV = \begin{pmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{pmatrix}\begin{pmatrix} -\frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{pmatrix} = \begin{pmatrix} \sqrt{10} & 0 \\ -2\sqrt{10} & 0 \\ -2\sqrt{10} & 0 \end{pmatrix}$$

13. *Suppose that the $m \times n$ matrix A has the singular value decomposition $A = U\Sigma V^T$ where U is an $m \times m$ orthogonal matrix, Σ is an $m \times n$ "diagonal" matrix with r positive entries and no negative entries, and V is an $n \times n$ orthogonal matrix. Show that the columns of V are eigenvectors of A^TA, the columns of U are eigenvectors of AA^T and that the diagonal entries of Σ are the singular values of A. (Text problem 125[19].)*

Viewing Σ as a matrix with $r \times r$ diagonal matrix block $D = \text{diag}(\sigma_1, \ldots, \sigma_r)$ in the upper left corner and zeros elsewhere we see that

$$\Sigma^T\Sigma = \begin{pmatrix} D^T_{r\times r} & 0^T_{(m-r)\times r} \\ 0^T_{r\times(n-r)} & 0^T_{(m-r)\times(n-r)} \end{pmatrix} \begin{pmatrix} D_{r\times r} & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix}$$

$$= \begin{pmatrix} D^2_{r\times r} & 0_{r\times(n-r)} \\ 0_{(n-r)\times r} & 0_{(n-r)\times(n-r)} \end{pmatrix}$$

$$\Sigma\Sigma^T = \begin{pmatrix} D_{r\times r} & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix} \begin{pmatrix} D^T_{r\times r} & 0^T_{(m-r)\times r} \\ 0^T_{r\times(n-r)} & 0^T_{(m-r)\times(n-r)} \end{pmatrix}$$

$$= \begin{pmatrix} D^2_{r\times r} & 0_{r\times(m-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(m-r)} \end{pmatrix}$$

where $D^2 = \text{diag}(\sigma_1,^2 \ldots, \sigma_r^2)$. We have $A = U\Sigma V^T$ so that the matrix A^TA acting on the columns of V yields

$$A^TAV = (U\Sigma V^T)^T(U\Sigma V^T)V = V\Sigma^TU^TU\Sigma V^TV = V\Sigma^T\Sigma = [\sigma_1^2\mathbf{v}_1, \ldots, \sigma_r^2\mathbf{v}_r, \mathbf{0}, \ldots, \mathbf{0}]$$

In other words, the $A^TA\mathbf{V}_i = \sigma^2\mathbf{v}_i$ for $i = 1, \ldots, n$, where $\sigma_1 = 0$ if $i > r$. This says that the $\mathbf{v}_i$ are eigenvectors whose eigenvalues $\lambda_i = \sigma_i^2$, namely, the σ_i's are singular values of A. To check that the columns of U are eigenvectors of AA^T we compute

$$AA^TU = (U\Sigma V^T)(U\Sigma V^T)^TU = U\Sigma V^TV\Sigma^TU^TU = U\Sigma\Sigma^T = [\sigma_1^2\mathbf{u}_1, \ldots, \sigma_r^2\mathbf{u}_r, \mathbf{0}, \ldots, \mathbf{0}]$$

In other words, the $AA^T\mathbf{u}_i = \sigma^2\mathbf{u}_i$ for $i = 1, \ldots, m$, where $\sigma_1 = 0$ is $i > r$. This says that the $\mathbf{u}_i$ are eigenvectors with eigenvalues $\lambda_i = \sigma_i^2$.

14. *Find the SVD of A. Hint: work with A^T. (Text problem 425[13].)*

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

Note that if we find SVD for $A^T = U\Sigma V^T$ then automatically we have the SVD for

$$A = (A^T)^T = (U\Sigma V^T)^T = V\Sigma^TU^T$$

where U and V reverse roles. Now

$$AA^T = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix}$$

The characteristic equation is

$$\begin{vmatrix} 17-\lambda & 8 \\ 8 & 17-\lambda \end{vmatrix} = (\lambda - 17)^2 - 8^2 = (\lambda - 9)(\lambda - 25)$$

The eigenvalues are $\lambda_1 = 25$ and $\lambda - 2 = 9$ so that the singular values are $\sigma_1 = 5$ and $\sigma_2 = 3$. Then we get the 3×2 "diagonal" matrix

$$\Sigma = \begin{pmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{pmatrix}$$

We get eigenvectors by inspection

$$(A - \lambda_1 I)\widetilde{\mathbf{v}_1} = \begin{pmatrix} -8 & 8 \\ 8 & -8 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad (A - \lambda_2 I)\widetilde{\mathbf{v}_2} = \begin{pmatrix} 8 & 8 \\ 8 & 8 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Normalizing, we find the orthogonal matrix

$$V = [\mathbf{v}_1, \mathbf{v}_2] = \left[\frac{\widetilde{\mathbf{v}_1}}{\|\widetilde{\mathbf{v}_1}\|}, \frac{\widetilde{\mathbf{v}_2}}{\|\widetilde{\mathbf{v}_2}\|} \right] = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

To construct U we find

$$A^T \mathbf{v}_1 = \begin{pmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{5}{\sqrt{2}} \\ \frac{5}{\sqrt{2}} \\ 0 \end{pmatrix}, \qquad A^T \mathbf{v}_2 = \begin{pmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ -2\sqrt{2} \end{pmatrix}$$

The last column has to be orthogonal to both $A^T\mathbf{v}_1$ and $A^T\mathbf{v}_2$. Thus it satifies the system

$$\begin{pmatrix} \frac{5}{\sqrt{2}} & \frac{5}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & -2\sqrt{2} \end{pmatrix} \mathbf{y} = \mathbf{0}; \qquad \mathbf{y} = \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}$$

which we find by inspection. By normalizing we obtain

$$U = \left[\frac{A^T\mathbf{v}_1}{\|A^T\mathbf{v}_1\|}, \frac{A^T\mathbf{v}_2}{\|A^T\mathbf{v}_2\|}, \frac{\mathbf{y}}{\|\mathbf{y}\|} \right] = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & -\frac{2\sqrt{2}}{3} & \frac{1}{3} \end{pmatrix}$$

By construction we have $A^T = U\Sigma V^T$. Thus we have our SVD for $A = V\Sigma^T U^T$. We check

$$AU = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & -\frac{4}{3\sqrt{2}} & \frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{5}{\sqrt{2}} & -\frac{3}{\sqrt{2}} & 0 \\ \frac{5}{\sqrt{2}} & \frac{3}{\sqrt{2}} & 0 \end{pmatrix}$$

$$V\Sigma^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} = \begin{pmatrix} \frac{5}{\sqrt{2}} & -\frac{3}{\sqrt{2}} & 0 \\ \frac{5}{\sqrt{2}} & \frac{3}{\sqrt{2}} & 0 \end{pmatrix}$$

.

15. *Use the pseudoinverse to find the shortest least squares solution of* $A\mathbf{x} = \mathbf{y}$. *(Like text problem 368[15] using the method of problem 434[15].)*

$$A = \begin{pmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 7 \\ 3 \\ 1 \end{pmatrix}$$

We found the SVD for $A = U\Sigma V^T$ in problem (10).

$$U\Sigma V^T = \begin{pmatrix} \frac{1}{3} & \frac{2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} \\ -\frac{2}{3} & \frac{1}{\sqrt{5}} & -\frac{4}{3\sqrt{5}} \\ -\frac{2}{3} & 0 & \frac{\sqrt{5}}{3} \end{pmatrix} \begin{pmatrix} 3\sqrt{10} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -\frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{pmatrix}$$

The rank of A is $r = 1$. The reduced SVD is thus

$$A = U_r D V_r^T = \begin{pmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ -\frac{2}{3} \end{pmatrix} \begin{pmatrix} 3\sqrt{10} \end{pmatrix} \begin{pmatrix} -\frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix}$$

The pseudoinverse is

$$A^+ = V_r D^{-1} U_r^T = \begin{pmatrix} -\frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} \frac{1}{3\sqrt{10}} \end{pmatrix} \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} \end{pmatrix} = \begin{pmatrix} -\frac{1}{30} & \frac{1}{15} & \frac{1}{15} \\ \frac{1}{90} & -\frac{1}{45} & -\frac{1}{45} \end{pmatrix}$$

the least squares solution is thus

$$\widehat{\mathbf{x}} = A^+\mathbf{b} = \begin{pmatrix} -\frac{1}{30} & \frac{2}{30} & \frac{2}{30} \\ \frac{1}{90} & -\frac{2}{90} & -\frac{2}{90} \end{pmatrix} \begin{pmatrix} 7 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{30} \\ -\frac{1}{90} \end{pmatrix}$$

To check, the projection onto $\operatorname{Col} A = \operatorname{span}\{\mathbf{a}_2\}$ of $\mathbf{b}$ is

$$\widehat{\mathbf{b}} = \frac{\mathbf{b} \bullet \mathbf{a}_2}{\mathbf{a}_2 \bullet \mathbf{a}_2}\mathbf{a}_2 = \frac{-1}{9}\begin{pmatrix} 1 \\ -2 \\ -2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{9} \\ \frac{2}{9} \\ \frac{2}{9} \end{pmatrix}$$

On the other hand $\widehat{\mathbf{x}}$ solves $A\widehat{\mathbf{x}} = \widehat{\mathbf{b}}$

$$A\widehat{\mathbf{x}} = \begin{pmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{pmatrix}\begin{pmatrix} \frac{3}{90} \\ -\frac{1}{90} \end{pmatrix} = \begin{pmatrix} -\frac{1}{9} \\ \frac{2}{9} \\ \frac{2}{9} \end{pmatrix}$$