MATH 223, Linear Algebra
Fall 2007
Solutions to Assignment 9

1. Let $W = Span \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ 3 \\ 8 \end{pmatrix} \right\}$ be a subspace of $\mathcal{R}^4$. Find an orthonormal basis for each of W and $W^{\perp}$. Also find the orthogonal projections $Proj_W \vec{v}$ and $Proj_{W^{\perp}} \vec{v}$, where $\vec{v} = \begin{pmatrix} 1 \\ 3 \\ 5 \\ 7 \end{pmatrix}$.

Solution: Starting the Gram-Schmidt process, let $\vec{w}_1 = \vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix}$.

Next let

$$\vec{w}_2 = \vec{v}_2 - \frac{< \vec{v}_2, \vec{w}_1 >}{< \vec{w}_1, \vec{w}_1 >} \vec{w}_1.$$

Here $\vec{v}_2 = \begin{pmatrix} 3 \\ 4 \\ 3 \\ 8 \end{pmatrix}$. So $< \vec{v}_2, \vec{w}_1 >= 1 \cdot 3 + 2 \cdot 4 + 1 \cdot 3 + 2 \cdot 8 = 30$ and $< \vec{w}_1, \vec{w}_1 >= 1^2 + 2^2 + 1^2 + 2^2 = 10$. So $\vec{w}_2 = \begin{pmatrix} 3 \\ 4 \\ 3 \\ 8 \end{pmatrix} - 3 \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ 0 \\ 2 \end{pmatrix}$. Note that this is indeed orthogonal to $\vec{v}_1 = \vec{w}_1$. The norm $||\vec{w}_1||$ of $\vec{w}_1$ is $\sqrt{10}$ and the norm $||\vec{w}_2||$ of $\vec{w}_2$ is $\sqrt{(-2)^2 + 0^2 + 2^2 + 0^2} = \sqrt{8} = 2\sqrt{2}$. An orthonormal basis for W is then

$$\left\{ \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix}, \frac{1}{2\sqrt{2}} \begin{pmatrix} 0 \\ -2 \\ 0 \\ 2 \end{pmatrix} \right\}.$$

(Note that the last vector can be simplified by cancelling 2's.)

We could continue with Gram-Schmidt to find an orthonormal basis for $W^{\perp}$, but it's likely easier to first solve $A\vec{x} = \vec{0}$, where $A = \begin{pmatrix} 1 & 2 & 1 & 2 \\ 0 & -2 & 0 & 2 \end{pmatrix}$.

This reduces to $\begin{pmatrix} 1 & 0 & 1 & 4 \\ 0 & 1 & 0 & -1 \end{pmatrix}$. A basis for its null space, which is $W^\perp$ is then $\left\{ \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}$. We turn this into an orthogonal basis with a single step of the Gram-Schmidt process. We replace the second vector $\vec{u}_2$ by $\vec{u}_2 - \frac{<\vec{u}_2,\vec{u}_1>}{<\vec{u}_1,\vec{u}_1>}\vec{u}_1$. ($\vec{u}_1$ is the first of these two vectors.) We get $\begin{pmatrix} -2 \\ 1 \\ -2 \\ 1 \end{pmatrix}$. Normalizing these two orthogonal vectors gives us an orthonormal basis for $W^\perp$. It is

$$\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{10}} \begin{pmatrix} -2 \\ 1 \\ -2 \\ 1 \end{pmatrix} \right\}.$$

$$Proj_W \vec{v} = \frac{< \vec{v}, \vec{w}_1 >}{< \vec{w}_1, \vec{w}_1 >}\vec{w}_1 + \frac{< \vec{v}, \vec{w}_2 >}{< \vec{w}_2, \vec{w}_2 >}\vec{w}_2 = \frac{26}{10}\begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix} + \frac{8}{8}\begin{pmatrix} 0 \\ -2 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{13}{5} \\ \frac{16}{5} \\ \frac{13}{5} \\ \frac{36}{5} \end{pmatrix}.$$

[It's important here that $\vec{w}_1$ and $\vec{w}_2$ are orthogonal; using the original $\vec{v}_1$ and $\vec{v}_2$ would not work.]

We could do something similar to find $Proj_{W^\perp}\vec{v}$, but it has to be

$$\vec{v} - Proj_W \vec{v} = \begin{pmatrix} -\frac{8}{5} \\ -\frac{1}{5} \\ \frac{12}{5} \\ -\frac{1}{5} \end{pmatrix}.$$

2. Let $W = Span \left\{ \begin{pmatrix} 1 \\ 1+i \\ 1 \\ 1-i \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 1+3i \\ 2i \\ 3-i \\ 2+i \end{pmatrix}, \begin{pmatrix} 2+7i \\ -1+5i \\ -2-i \\ 5+i \\ -1+3i \end{pmatrix} \right\}$ be a subspace of $\mathcal{C}^5$. Find an orthonormal basis for each of W and $W^\perp$. Also find the orthogonal projections $Proj_W \vec{v}$ and $Proj_{W^\perp}\vec{v}$, where $\vec{v} = \begin{pmatrix} 2-i \\ 2+2i \\ -2+i \\ 2-2i \\ -1 \end{pmatrix}$.

Solution: Let $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$ be (in order) the vectors in the given basis for

W. We let $\vec{w}_1 = \vec{v}_1$ and

$$\vec{w}_2 = \vec{v}_2 - \frac{< \vec{v}_2, \vec{w}_1 >}{< \vec{w}_1, \vec{w}_1 >} \vec{w}_1.$$

Now $< \vec{v}_2, \vec{w}_1 >= 4(1) + (1+3i)(1-i) + 2i(1) + (3-i)(1+i) + (2+i)(1) = 14 + 7i$. [I trust you see exactly why just those sign changes were made.] $< \vec{w}_1, \vec{w}_1 >= 1(1) + (1+i)(1-i) + 1(1) + (1-i)(1+i) + 1(1) = 7$. $\vec{w}_2 = \vec{v}_2 - (2+i)\vec{w}_1 = \begin{pmatrix} 2-i \\ 0 \\ -2+i \\ 0 \\ 0 \end{pmatrix}$. To find $\vec{w}_3$, we compute

$$\vec{v}_3 - \frac{< \vec{v}_3, \vec{w}_1 >}{< \vec{w}_1, \vec{w}_1 >} \vec{w}_1 - \frac{< \vec{v}_3, \vec{w}_2 >}{< \vec{w}_2, \vec{w}_2 >} \vec{w}_2.$$

[It would be legitimate to use a scalar multiple of $\vec{w}_2$ instead of $\vec{w}_2$ here, but *NOT* the original $\vec{v}_2$.] $< \vec{v}_3, \vec{w}_1 >= 7 + 21i$, $< \vec{w}_1, \vec{w}_1 >= 7$, $< \vec{v}_3, \vec{w}_2 >= 20i$ and $< \vec{w}_2, \vec{w}_2 >= 10$. So $\vec{w}_3 = \vec{v}_3 - (1+3i)\vec{w}_1 - 2i\vec{w}_2 = \begin{pmatrix} -1 \\ 1+i \\ -1 \\ 1-i \\ -2 \end{pmatrix}$. We now have an orthogonal basis for W; normalizing it gives us the orthonormal basis (for W)

$$\left\{ \frac{1}{\sqrt{7}} \begin{pmatrix} 1 \\ 1+i \\ 1 \\ 1-i \\ 1 \end{pmatrix}, \frac{1}{\sqrt{10}} \begin{pmatrix} 2-i \\ 0 \\ -2+i \\ 0 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{10}} \begin{pmatrix} -1 \\ 1+i \\ -1 \\ 1-i \\ -2 \end{pmatrix} \right\}.$$

To find a basis for $W^{\perp}$, we start by solving $A\vec{x} = \vec{0}$, where $A = \begin{pmatrix} 1 & 1+i & 1 & 1-i & 1 \\ 2-i & 0 & -2+i & 0 & 0 \\ -1 & 1+i & -1 & 1-i & -2 \end{pmatrix}$. A row-reduces to $\begin{pmatrix} 1 & 0 & 0 & 0 & \frac{3}{4} \\ 0 & 1 & 0 & -i & -\frac{1}{4} + \frac{1}{4}i \\ 0 & 0 & 1 & 0 & \frac{3}{4} \end{pmatrix}$. A basis for the null space of A is then (avoiding fractions as long as possible)

$$\left\{ \begin{pmatrix} 0 \\ i \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 1-i \\ -3 \\ 0 \\ 4 \end{pmatrix} \right\}.$$

But these vectors are not in $W^\perp$, since the dot product of each of them with the vectors in W is zero — we want that the dot product of their conjugates with the vectors in W is 0. This is easy to fix — take conjugates, and we have a basis for $W^\perp$; it is

$$\left\{ \begin{pmatrix} 0 \\ -i \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 1+i \\ -3 \\ 0 \\ 4 \end{pmatrix} \right\}.$$

We aren't done yet; we need to find an orthonormal basis for $W^\perp$. To this end, we call these two vectors $\vec{u}_1$ and $\vec{u}_2$ and replace $\vec{u}_2$ by $\vec{u}_2 - \frac{<\vec{u}_2, \vec{u}_1>}{<\vec{u}_1, \vec{u}_1>}\vec{u}_1 = \vec{u}_2 - \frac{-1+i}{2}\vec{u}_1 = \begin{pmatrix} -3 \\ \frac{1+i}{2} \\ -3 \\ \frac{1-i}{2} \\ 4 \end{pmatrix}$. Now we normalize these two vectors and get our orthonormal basis for $W^\perp$. It is

$$\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -i \\ 0 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{140}} \begin{pmatrix} -6 \\ 1+i \\ -6 \\ 1-i \\ 8 \end{pmatrix} \right\}.$$

To find the projection $Proj_W \vec{v}$ using Gram-Schmidt, we need to use an orthogonal basis of W; the original basis $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ will not work, but $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ will. The projection is

$$\frac{< \vec{v}, \vec{w}_1 >}{< \vec{w}_1, \vec{w}_1 >}\vec{w}_1 + \frac{< \vec{v}, \vec{w}_2 >}{< \vec{w}_2, \vec{w}_2 >}\vec{w}_2 + \frac{< \vec{v}, \vec{w}_3 >}{< \vec{w}_3, \vec{w}_3 >}\vec{w}_3,$$

and this is $\frac{7}{7}\vec{w}_1 + \frac{10}{10}\vec{w}_2 + \frac{10}{10}\vec{w}_3 = \begin{pmatrix} 2-i \\ 2+2i \\ -2+i \\ 2-2i \\ -1 \end{pmatrix}$.

Finally, $Proj_{W^\perp}\vec{v} = \vec{v} - Proj_W \vec{v} = \vec{0}$. This happens, of course, because $\vec{v} \in W$. I realize that there are a lot of calculations in this problem; I trust you noticed that I went out of my way to keep them relatively fraction-free. Between this one and problem 1, you should appreciate all the complications that can can arise in this kind of calculation (and at least some of the shortcuts).

3. For each of the following Hermitian matrices H_j, find a unitary matrix U_j such that $\overline{U}_j^T H_j U_j$ is diagonal. Also find the diagonal matrix.

$H_1 = \begin{pmatrix} 2 & 1+i \\ 1-i & 3 \end{pmatrix}$, $H_2 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$, $H_3 = \begin{pmatrix} 4 & 2+2i & 1-i \\ 2-2i & 6 & -2i \\ 1+i & 2i & 3 \end{pmatrix}$.

Solution: The characteristic polynomial for H_1 is $(\lambda-2)(\lambda-3)-(-1-i)(-1+i) = \lambda^2-5\lambda+4 = (\lambda-4)(\lambda-1)$. $4I-H_1 = \begin{pmatrix} 2 & -1-i \\ -1+i & -1 \end{pmatrix}$; an eigenvector is $\begin{pmatrix} 1 \\ 1-i \end{pmatrix}$. $I-H_1 = \begin{pmatrix} -1 & -1-i \\ -1+i & -2 \end{pmatrix}$; an e-vector for 1 is $\begin{pmatrix} -1-i \\ 1 \end{pmatrix}$. You will note that these e-vectors are orthogonal (as they must be) but not yet normalized; the norm of each is $\sqrt{3}$, so $U_1 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & -1-i \\ 1-i & 1 \end{pmatrix}$ is unitary, and $U_1^{-1}H_1U_1 = \overline{U}_1^T H_1 U_1 = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$. [This is not the only possible answer.]

I hope at least some of you recognize by sight (by now) that 1 is an e-value of the matrix H_2 — of multiplicity two — and that the only other e-value has to be 4 (as the trace of H_2 is $6 = 1+1+4$). But anyway, the usual calculation gives that $det(\lambda I - H_2) = \lambda^3 - 6\lambda^2 + 9\lambda - 4 = (\lambda-1)^2(\lambda-4)$. $I - H_2 = \begin{pmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix}$. This row-reduces almost immediately to $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and it's easy to find two independent e-vectors, say $\vec{v}_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ and $\vec{v}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$; these are not orthogonal, so we replace the second one by $\vec{v}_2 - \frac{<\vec{v}_2,\vec{v}_1>}{<\vec{v}_1,\vec{v}_1>}\vec{v}_1 = \begin{pmatrix} -\frac{1}{2} \\ 1 \\ -\frac{1}{2} \end{pmatrix}$. Normalizing these two vectors gives us the first two columns of U_2. They are $\begin{pmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ and $\begin{pmatrix} -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \end{pmatrix}$. $4I - H_2 = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$, which row-reduces to $\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$. A normalized eigenvector is $\frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. A possible unitary (in fact, orthogonal) U_2 such that $U_2^T H_2 U_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ is

$U_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}$.

Here we skip some of the details. The characteristic polynomial of H_3 is $\lambda^3 - 13\lambda^2 + 40\lambda - 36 = (\lambda-2)^2(\lambda-9)$. $2I - H_3$ row-reduces to $\begin{pmatrix} 1 & 1+i & \frac{1}{2} - \frac{1}{2}i \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. $\begin{pmatrix} -1-i \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} -1+i \\ 0 \\ 2 \end{pmatrix}$ are e-vectors, but not orthogonal; we replace the second one using Gram-Schmidt to get $\begin{pmatrix} -\frac{1}{3} + \frac{1}{3}i \\ \frac{2}{3}i \\ 2 \end{pmatrix}$. Normalized versions will show up in U_3. $9I - H_3$ row-reduces to $\begin{pmatrix} 1 & 0 & -1+i \\ 0 & 1 & 2i \\ 0 & 0 & 0 \end{pmatrix}$; an e-vector is $\begin{pmatrix} 1-i \\ -2i \\ 1 \end{pmatrix}$. Normalizing the three orthogonal eigenvectors, we get $U_3 = \begin{pmatrix} -\frac{1}{\sqrt{3}} - \frac{1}{\sqrt{3}}i & -\frac{1}{\sqrt{42}} + \frac{1}{\sqrt{42}}i & \frac{1}{\sqrt{7}} - \frac{1}{\sqrt{7}}i \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{42}}i & -\frac{2}{\sqrt{7}}i \\ 0 & \frac{6}{\sqrt{42}} & \frac{1}{\sqrt{7}} \end{pmatrix}$. $\overline{U}_3^T H_3 U_3 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 9 \end{pmatrix}$.

4. (a) Suppose that A is a symmetric matrix over the reals with nonnegative eigenvalues. Show that there is a symmetric real matrix B such that $B^2 = A$. [Such a matrix is called a *square root* of A.]

Solution: We know that there is an orthogonal matrix P such that $P^T AP = D$, where D is diagonal, and its diagonal entries are the eigenvalues of A. We are assuming these are nonnegative, so each has at least one real square root (usually two). Also, as $P^T = P^{-1}$, $A = PDP^T$. Let C be a diagonal matrix where we replace each diagonal entry of D by one of its square roots. Let $B = PCP^T$ — note that B is symmetric since $B^T = (P^T CP)^T = P^T C^T (P^T)^T$ and $C = C^T$ as C is diagonal. Now $B^2 = (PCP^T)(PCP^T) = PC^2P^T = PDP^T = A$. [Here we use that C is diagonal, so $C^2 = D$; we also use again that $P^T = P^{-1}$ as P is orthogonal.]

(b) Find a symmetric square root of the matrix $\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$.

Solution: We did most of the work already. With $A = H_2$ and $P = U_2$ from problem 3, we have $P^T AP = D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$. Let $C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, so $C^2 = D$. We let $B = PCP^T = \frac{1}{3} \begin{pmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{pmatrix}$, and we have that $B^2 = A$. (This is easy to check.)

5. Suppose that (a_1, b_1),...,(a_n, b_n) are points in the plane $\mathcal{R}^2$. If $n \geq 3$, it's unlikely that there is a line going through all of them.

(a) Show that the line defined by $y = \alpha x + \beta$ goes through all of them if and only if $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is a solution to $A\vec{x} = \vec{b}$, where $A = \begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix}$ and $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$.

Solution: This is just a restatement. To say that $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ solves $A\vec{x} = \vec{b}$ is the same thing as saying that $a_j\alpha + \beta = b_j$ for each $j = 1, \ldots, n$. But that's no different from saying that each point (a_j, b_j) is on the line defined by $y = \alpha x + \beta$.

As mentioned, this is usually impossible if $n \geq 3$. But the system $A^T A\vec{x} = A^T\vec{b}$ frequently has a unique solution.

(b) Show that, if A is as above, and not all the values a_1,...,a_n are the same, then $A^T A$ has positive determinant and hence there is a unique solution to $A^T A\vec{x} = A^T\vec{b}$.

Solution: Suppose that $\vec{a}$ is the first column of A and $\vec{1}$ the second column. Then $\vec{a}^T$ and $\vec{1}^T$ are the rows of A^T. $A^T A$ is the symmetric 2×2 matrix $\begin{pmatrix} \vec{a} \cdot \vec{a} & \vec{a} \cdot \vec{1} \\ \vec{1} \cdot \vec{a} & \vec{1} \cdot \vec{1} \end{pmatrix}$.

[Oh, $\vec{1} \cdot \vec{1} = n$, but that's not terribly important here.]

The determinant of $A^T A$ is then $(\vec{a} \cdot \vec{a})(\vec{1} \cdot \vec{1}) - (\vec{a} \cdot \vec{1})^2$. This is always nonnegative by Cauchy-Schwartz, and the same theorem guarantees that it's strictly greater than 0 unless $\vec{a}$ is a scalar multiple of $\vec{1}$.

That is $det(A^T A) > 0$ unless $a_1 = a_2 = \ldots = a_n$. The system $(A^T A)\vec{x} = A^T\vec{b}$ then has the unique solution $\vec{x} = (A^T A)^{-1}A^T\vec{b}$.

[Incidentally, in case all the a_j's are the same, all the experimental data will be on single vertical straight line — this is extremely unlikely.]

The solution $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is known as the *least squares solution* to the original system $A\vec{x} = \vec{b}$, and the line defined by $y = \alpha x + \beta$ is the *line of best fit* (to the data (a_1, b_1),. . . ,(a_n, b_n)).

[The reason for this terminology is more or less as follows: If we want to mimic our experimental data with a simple mathematical model, we might "pretend" that it follows a simple formula, like $y = \alpha x + \beta$; real life is rarely that simple, but sometimes it comes close. That is, sometimes it reasonable to assume — on the basis of the data you know — that a straight line model is approximately right. If the given straight line defined by $y = \alpha x + \beta$ were a perfect model, then $\alpha a_j + \beta = b_j$ for every j; if not, then $\alpha a_j + \beta = c_j$ for some c_j, which we hope is not too far from b_j. If we trade in the α and β we computed above for some other values α^* and β^*, we'd get the line $y = \alpha^* x + \beta^*$ and instead of $\vec{c}$, we'd get $\vec{c}^*$ with $c_j^* = \alpha^* a_j + \beta^*$. $\vec{c}^*$ will be in the space W spanned by $\vec{a}$ and $\vec{1}$. The closest vector in this space to $\vec{b}$ is $Proj_W\vec{b}$ and it can be shown that this is the $\vec{c}$ described above. That is, $||\vec{b} - \vec{c}|| < ||\vec{b} - \vec{c}^*||$ for any such $\vec{c}^*$.]

(c) If $n = 4$ and $(a_1, b_1) = (0, 2)$, $(a_2, b_2) = (1, 3)$, $(a_3, b_3) = (2, 5)$ and $(a_4, b_4) = (3, 6)$ find the least-squares solution to $A\vec{x} = \vec{b}$, and the line of best fit to the data (a_1, b_1), (a_2, b_2), (a_3, b_3) and (a_4, b_4).

Solution: $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}$ and $\vec{b} = \begin{pmatrix} 2 \\ 3 \\ 5 \\ 6 \end{pmatrix}$. It's easy to see that $A\vec{x} = \vec{b}$ has no solution. But $A^T A = \begin{pmatrix} 14 & 6 \\ 6 & 4 \end{pmatrix}$ and $A^T\vec{b} = \begin{pmatrix} 31 \\ 16 \end{pmatrix}$. The unique solution to $(A^T A)\vec{x} = A^T\vec{b}$ is $\vec{x} = \begin{pmatrix} 1.4 \\ 1.9 \end{pmatrix}$. That's the least squares solution; the line of best fit is defined by $y = 1.4x + 1.9$.

[The method of least squares is frequently used in certain applications of mathematics and statistics. If the data suggests some kind of model other than a straight line (say a quadratic, or more complicated curve) the method can be easily adapted.]