

METHODS OF APPLIED MATH
Fall 2009

Assignment 2: Solution Outline

1. Consider two sets of rectangular cartesian coordinate systems for \mathcal{R}^3 with a common origin. Unit vectors in the directions of the three coordinate axes for the two coordinate systems serve as two different orthonormal bases. In terms of the cosines of the angles between the two sets of axes, find the matrix \mathbf{A} relating the coordinates of a vector in the two coordinate systems and show that \mathbf{A} is orthogonal.

Let the two sets of rectangular axes be defined by the orthonormal sets of basis vectors $E : \{\mathbf{e}_i\}$ and $E' : \{\mathbf{e}'_i\}$. Let us suppose that α_{1j} , α_{2j} and α_{3j} are the angles that the unit vector \mathbf{e}_j makes with the unit vectors \mathbf{e}'_1 , \mathbf{e}'_2 and \mathbf{e}'_3 , respectively. Then we can express \mathbf{e}_j as a linear combination of the \mathbf{e}'_i as follows.

$$\mathbf{e}_j = \cos \alpha_{1j} \mathbf{e}'_1 + \cos \alpha_{2j} \mathbf{e}'_2 + \cos \alpha_{3j} \mathbf{e}'_3.$$

On using the notation

$$a_{ij} \equiv \cos \alpha_{ij},$$

the above expression takes the form

$$\mathbf{e}_j = a_{ij} \mathbf{e}'_i.$$

Recall that this is the familiar connection between bases we have been using all along. The above discussion merely points out that the entries a_{ij} of the transformation matrix \mathbf{A} are simply the cosines of the angles between the two sets of axes.

In order to show that \mathbf{A} is orthogonal, i.e., its columns are pairwise orthogonal unit vectors, let us denote the components of a vector \mathbf{x} by x_i and x'_i in the systems E and E' respectively. Then we have

$$\mathbf{x} = x_j \mathbf{e}_j = x_j a_{ij} \mathbf{e}'_i = a_{ij} x_j \mathbf{e}'_i = x'_i \mathbf{e}'_i,$$

so that

$$x'_i = a_{ij} x_j.$$

Now we know that the norm of \mathbf{x} should be the same in both coordinate systems. Therefore,

$$\begin{aligned} \|\mathbf{x}\|^2 &= x'_i x'_i \\ &= a_{ij} x_j a_{ik} x_k \\ &= a_{ij} a_{ik} x_j x_k. \end{aligned}$$

Also,

$$\begin{aligned} \|\mathbf{x}\|^2 &= x_j x_j \\ &= x_j x_k \delta_{jk}. \end{aligned}$$

Upon equating the two different expressions for $\|\mathbf{x}\|^2$ we get

$$[a_{ij} a_{ik} - \delta_{jk}] x_j x_k = 0.$$

For this relation to hold for arbitrary \mathbf{x} , one must have

$$a_{ij} a_{ik} = \delta_{jk},$$

i.e., the columns of \mathbf{A} are pairwise orthogonal unit vectors.

2. (a) Let V be a linear vector space and W a subspace of V . Let \mathbf{y} be a vector in W , \mathbf{x} a vector in V but not in W , and \mathbf{w} the projection of \mathbf{x} onto W . Show that $\|\mathbf{x} - \mathbf{w}\| \leq \|\mathbf{x} - \mathbf{y}\|$.

(b) In R^3 , find the projection of $\begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$ onto the plane $2x_1 + x_2 - 3x_3 = 0$. Show that the length of the vector difference between the vector and its projection is equal to the distance of the vector from the plane.

(c) Find all vectors in R^4 which are orthogonal to the subspace defined by the equations $x_1 + x_2 = 0$, $x_4 - x_3 = 0$.

(a) Since $\mathbf{x} - \mathbf{y} = \mathbf{x} - \mathbf{w} + \mathbf{w} - \mathbf{y}$,

$$\begin{aligned} \|\mathbf{x} - \mathbf{y}\|^2 &= ((\mathbf{x} - \mathbf{w}) + (\mathbf{w} - \mathbf{y}), (\mathbf{x} - \mathbf{w}) + (\mathbf{w} - \mathbf{y})) \\ &= (\mathbf{x} - \mathbf{w}, \mathbf{x} - \mathbf{w}) + (\mathbf{w} - \mathbf{y}, \mathbf{w} - \mathbf{y}) + (\mathbf{x} - \mathbf{w}, \mathbf{w} - \mathbf{y}) + (\mathbf{w} - \mathbf{y}, \mathbf{x} - \mathbf{w}). \end{aligned}$$

As \mathbf{w} is the orthogonal projection of \mathbf{x} onto W , $\mathbf{x} - \mathbf{w}$ is orthogonal to all vectors in W , in particular to $\mathbf{w} - \mathbf{y} \in W$. Therefore the last two terms in the RHS of the last expression above vanish, yielding

$$\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x} - \mathbf{w}\|^2 + \|\mathbf{w} - \mathbf{y}\|^2.$$

Since $\|\mathbf{w} - \mathbf{y}\| \geq 0$, $\|\mathbf{x} - \mathbf{w}\| \leq \|\mathbf{x} - \mathbf{y}\|$.

(b) The plane defined by $2x_1 + x_2 - 3x_3 = 0$ is spanned by the independent vectors

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} \text{ and } \mathbf{u}_2 = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}.$$

Upon using Gram-Schmidt orthogonalization, we find an equivalent orthonormal basis for the plane, $\mathbf{v}_1, \mathbf{v}_2$, computed as follows.

$$\mathbf{v}_1 = \frac{1}{\|\mathbf{u}_1\|} \mathbf{u}_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix},$$

and

$$\begin{aligned} \hat{\mathbf{v}}_2 &= \mathbf{u}_2 - (\mathbf{u}_2, \mathbf{v}_1) \mathbf{v}_1 \\ &= \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix} - \frac{(-6)}{\sqrt{5}} \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 6 \\ 3 \\ 5 \end{bmatrix} \text{ leading to} \\ \mathbf{v}_2 &= \frac{1}{\|\hat{\mathbf{v}}_2\|} \hat{\mathbf{v}}_2 = \frac{1}{\sqrt{70}} \begin{bmatrix} 6 \\ 3 \\ 5 \end{bmatrix}. \end{aligned}$$

Now the projection of $\mathbf{x} = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$ is given by

$$\begin{aligned} \mathbf{w} &= (\mathbf{x}, \mathbf{v}_1) \mathbf{v}_1 + (\mathbf{x}, \mathbf{v}_2) \mathbf{v}_2 \\ &= \frac{1}{14} \begin{bmatrix} 52 \\ 61 \\ 55 \end{bmatrix}. \end{aligned}$$

Now,

$$\mathbf{d} = \mathbf{x} - \mathbf{w} = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} - \frac{1}{14} \begin{bmatrix} 52 \\ 61 \\ 55 \end{bmatrix} = \frac{5}{14} \begin{bmatrix} -2 \\ -1 \\ 3 \end{bmatrix}.$$

Then, the magnitude of the vector difference between \mathbf{x} and its projection \mathbf{w} is given by

$$\|\mathbf{d}\| = \frac{5}{\sqrt{14}}.$$

Note that the origin lies on the plane, and that the unit normal to the plane is

$$\mathbf{n} = \frac{1}{\sqrt{14}} \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}.$$

Therefore the distance of the vector \mathbf{x} from the plane is the component of \mathbf{x} along \mathbf{n} , i.e.,

$$|(\mathbf{x}, \mathbf{n})| = \frac{5}{\sqrt{14}}.$$

(c) A typical vector in the subspace W defined by the given equations is

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

Thus W has dimension 2, and it is spanned by the vectors

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Let the vector

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}$$

be orthogonal to W . Then it is orthogonal to both \mathbf{u}_1 and \mathbf{u}_2 , leading to the constraints

$$y_1 - y_2 = 0, \quad y_3 + y_4 = 0.$$

Then

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_1 \\ y_3 \\ -y_3 \end{bmatrix} = y_1 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + y_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix},$$

so that all vectors orthogonal to W themselves form a 2-D subspace of \mathcal{R}^4 .

3. (a) Find the best least-squares polynomial approximation of degree 2 for the function $f(x) = \cos(x/2)$ in the interval $[-\pi, \pi]$ using the inner product

$$\int_{-\pi}^{\pi} f(x)g(x) dx.$$

What is the \mathcal{L}^2 error of this approximation?

- (b) Compute the Taylor series $T(x)$ of $f(x)$ about $x = 0$ up to the x^2 term. On a graph plot $f(x)$ and its two approximations, $g(x)$ and $T(x)$, on the interval $[-\pi, \pi]$. Globally, which approximation do you prefer and why?

- (a) The desired polynomial is the orthogonal projection of f onto $\mathcal{P}^2[-\pi, \pi]$, the subspace of polynomials of degree 2, which is spanned by $p_0 = 1$, $p_1 = x$ and $p_2 = x^2$. The first task is to use the Gram-Schmidt procedure to derive an orthogonal basis for \mathcal{P}^2 , using the inner product given. Let the elements in this basis be $g_0(x)$, $g_1(x)$ and $g_2(x)$. Before proceeding further we compute the following quantities, which will be needed in the sequel.

$$\begin{aligned}\|p_0\| &= \sqrt{\int_{-\pi}^{\pi} 1 \, dx} = \sqrt{2\pi}, \\ \|p_1\| &= \sqrt{\int_{-\pi}^{\pi} x^2 \, dx} = \sqrt{\frac{2\pi^3}{3}}, \\ \|p_2\| &= \sqrt{\int_{-\pi}^{\pi} x^4 \, dx} = \sqrt{\frac{2\pi^5}{5}}.\end{aligned}$$

We begin computing the g_i by setting

$$g_0 = \frac{1}{\|p_0\|} p_0 = \frac{1}{\sqrt{2\pi}}.$$

Then,

$$\begin{aligned}\hat{g}_1 &= p_1 - (p_1, g_0)g_0 = x - \left(\int_{-\pi}^{\pi} \frac{x}{\sqrt{2\pi}} dx\right) \frac{1}{\sqrt{2\pi}} = x, \\ g_1 &= \frac{1}{\|\hat{g}_1\|} \hat{g}_1 = \sqrt{\frac{3}{2\pi^3}} x, \\ \hat{g}_2 &= p_2 - (p_2, g_0)g_0 - (p_2, g_1)g_1, \\ &= x^2 - \left(\int_{-\pi}^{\pi} \frac{x^2}{\sqrt{2\pi}} dx\right) \frac{1}{\sqrt{2\pi}} - \left(\int_{-\pi}^{\pi} \frac{3x^3}{\sqrt{2\pi^3}} dx\right) \frac{\sqrt{3}x}{\sqrt{2\pi^3}} \\ &= x^2 - \frac{\pi^2}{3}, \\ g_2 &= \frac{1}{\|\hat{g}_2\|} \hat{g}_2 = \sqrt{\frac{5}{8\pi^5}} (3x^2 - \pi^2).\end{aligned}$$

In the last step the following result has been used.

$$\|\hat{g}_2\| = \sqrt{\int_{-\pi}^{\pi} \left(x^2 - \frac{\pi^2}{3}\right)^2 dx} = \sqrt{\frac{8\pi^5}{45}}.$$

With the orthonormal basis computed, the desired polynomial approximation is given by

$$g(x) = (f, g_0)g_0 + (f, g_1)g_1 + (f, g_2)g_2.$$

The inner products appearing above are computed as

$$\begin{aligned}(f, g_0) &= \int_{-\pi}^{\pi} \frac{\cos(x/2)}{\sqrt{2\pi}} dx = \frac{4}{\sqrt{2\pi}}, \\ (f, g_1) &= \int_{-\pi}^{\pi} \frac{\sqrt{3}x \cos(x/2)}{\sqrt{2\pi^3}} dx = 0, \\ (f, g_2) &= \int_{-\pi}^{\pi} \sqrt{\frac{5}{8\pi^5}} (3x^2 - \pi^2) \cos(x/2) dx = \sqrt{\frac{40}{\pi^5}} (\pi^2 - 12).\end{aligned}$$

On substituting into the above expression for F we obtain

$$g(x) = \frac{2}{\pi} + \frac{5(\pi^2 - 12)(3x^2 - \pi^2)}{\pi^5}.$$

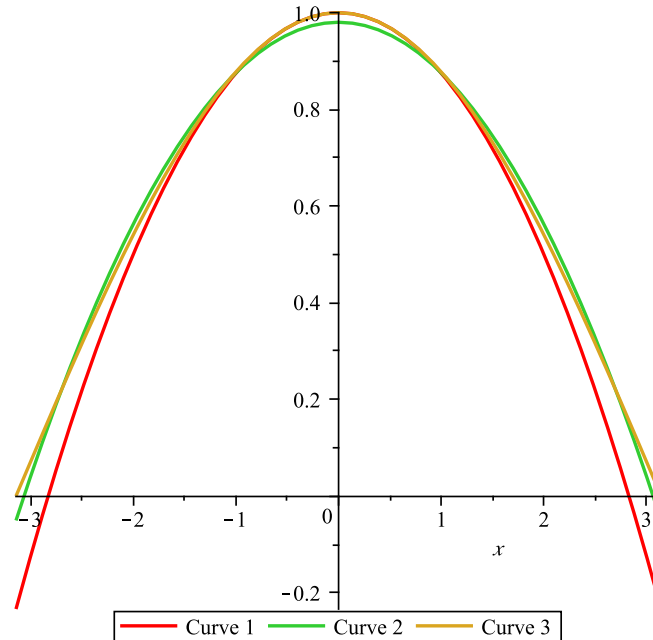


Figure 1: Graphs of $T(x)$ (curve 1), $g(x)$ (curve 2) and $f(x)$ (curve 3).

The \mathcal{L}^2 error is given by

$$E = \sqrt{\int_{-\pi}^{\pi} [\cos(x/2) - g(x)]^2 dx} = \sqrt{\frac{\pi^6 - 48\pi^4 + 960\pi^2 - 5760}{\pi^5}} = 0.043274.$$

(b) The 2-term Taylor series of $f = \cos(x/2)$ is

$$T(x) = 1 - \frac{x^2}{8}.$$

Figure 1 shows the graphs of g , T and f . Clearly, g is a better approximation globally, while T provides a better approximation near $x = 0$ but becomes progressively worse away from $x = 0$.

4. Consider the system with three springs and two weights shown in the figure. Following the discussion in class, write down the matrix equation governing the displacements x_1 , x_2 of the two weights W_1 , W_2 . For unit weights and spring constants, what are the displacements?

The extensions in the three springs are given by $e_1 = x_1$, $e_2 = x_2 - x_1$, $e_3 = x_2$. In matrix form,

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \text{or} \quad \mathbf{e} = \mathbf{A}\mathbf{x}.$$

The tensions in the springs are $T_1 = k_1 e_1$, $T_2 = k_2 e_2$, $T_3 = k_3 e_3$. In matrix form,

$$\begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} = \begin{bmatrix} k_1 & 0 & 0 \\ 0 & k_2 & 0 \\ 0 & 0 & k_3 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}, \quad \text{or} \quad \mathbf{T} = \mathbf{K}\mathbf{e}.$$

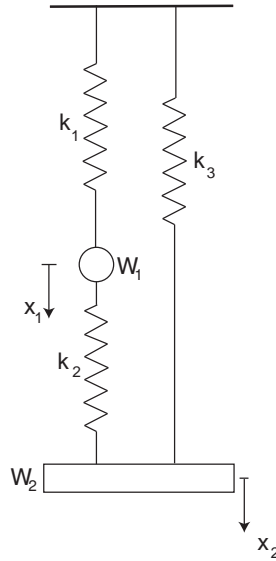


Figure 2: The spring-mass system.

The external forces due to the weights are balanced as $W_1 = T_1 - T_2$, $W_2 = T_2 + T_3$. In matrix form,

$$\begin{bmatrix} W_1 \\ W_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix}, \quad \text{or } \mathbf{W} = \mathbf{A}^T \mathbf{T}.$$

Combining these results, the displacements \mathbf{x} satisfy

$$\mathbf{A}^T \mathbf{K} \mathbf{A} \mathbf{x} = \mathbf{W},$$

or, upon matrix multiplication, and with $k_i = W_i = 1$, the problem reduces to

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

yielding $x_1 = x_2 = 1$.

5. (a) Determine whether the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}$$

is positive definite. Justify your answer. If the matrix is positive definite, then factorize it as $\mathbf{A} = \bar{\mathbf{L}}\bar{\mathbf{L}}^T$.

- (b) What symmetric matrix corresponds to the quadratic form $Q = x_1^2 + 10x_1x_2 + 25x_2^2$? Give an example of a nonsymmetric matrix, if there is one, that also corresponds to Q .
- (a) We carry out Gaussian elimination. In step 1, replace R_2 by $R_2 - R_1$ and R_3 by $R_3 - R_1$. Note that the pivot $d_1 = 1$ and the multipliers are $\ell_{12} = \ell_{31} = 1$. The result is

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

In step 2, replace R_3 by $R_3 - R_2$. The pivot is $d_2 = 1$ and the multiplier $\ell_{32} = 1$. the result is the upper triangular matrix

$$\mathbf{U} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

and the remaining pivot is $d_3 = 1$. Since all pivots are positive, \mathbf{A} is positive definite. Also, the lower triangular matrix \mathbf{L} formed by the multipliers is

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

The \mathbf{LU} decomposition

$$\mathbf{A} = \mathbf{LU}$$

is already in the desired form since in this case, $\mathbf{U} = \mathbf{L}^T$.

- (b) The symmetric matrix is

$$\begin{bmatrix} 1 & 5 \\ 5 & 25 \end{bmatrix}.$$

The non-symmetric matrix

$$\begin{bmatrix} 1 & 5+a \\ 5-a & 25 \end{bmatrix}.$$

will also yield the same quadratic form, for any a .

6. (a) Consider the overdetermined linear system $\mathbf{Ax} = \mathbf{b}$, where \mathbf{A} is an $m \times n$ matrix and $m > n$. While this system does not have a solution in general, it does so in the least squares sense. In other words, the solution \mathbf{x} minimizes the error $\|\mathbf{r}\|$ where $\mathbf{r} = \mathbf{Ax} - \mathbf{b}$ is the *residual* vector. Show that \mathbf{r} must be orthogonal to each column of \mathbf{A} . Use this fact to derive the *normal equations* $\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b}$, which we had derived in class by a different argument.
- (b) If a vertical beam has a downward force applied at its lower end, the beam stretches by an amount proportional to the applied force. The total length of the beam can therefore be modelled by the equation $y = x_1 + x_2 t$, where x_1 is its original length, t the applied force and x_2 the proportionality constant. Suppose that the following measurements are taken: $y = 11.60, 11.85, 12.25$ corresponding to $t = 10, 15, 20$ respectively.
- Set up the overdetermined 3×2 system of linear equations corresponding to the data collected.
 - Is the system consistent? If not, then compute each possible pair of values for x_1 and x_2 obtained by selecting any two of the equations of the system. Is there any reason to prefer any one of these results?
 - Set up the system of normal equations and solve it to obtain the least squares solution. Compare your results with those obtained in part (ii) above. Plot the data and the least-squares fit on a graph.
- (a) Let \mathbf{b}_c be a vector in the column space of \mathbf{A} . Then we can write $\mathbf{b} = \mathbf{b}_c + (\mathbf{b} - \mathbf{b}_c)$. In Problem 2(a) above we saw that $\|\mathbf{b} - \mathbf{b}_c\|$ is minimized over \mathbf{b}_c when \mathbf{b}_c is the orthogonal projection of \mathbf{b} onto the column space of \mathbf{A} . Then, $\mathbf{b} - \mathbf{b}_c$ is orthogonal to the column space of \mathbf{A} . With this choice of \mathbf{b}_c , let us choose \mathbf{x} such that $\mathbf{Ax} = \mathbf{b}_c$; that such an \mathbf{x} exists is guaranteed by \mathbf{b}_c being in the column space of \mathbf{A} . Now,

$$\mathbf{r} = \mathbf{Ax} - \mathbf{b} = (\mathbf{Ax} - \mathbf{b}_c) + (\mathbf{b}_c - \mathbf{b}) = \mathbf{b}_c - \mathbf{b}.$$

Hence the above choice of \mathbf{x} minimizes $\|\mathbf{r}\|$, and as shown above, \mathbf{r} is orthogonal to the column space of \mathbf{A} . This last result can be written in the index notation as $\{\mathbf{A}\}_{ij} r_i = 0$, or equivalently, as $\{\mathbf{A}\}_{ji}^T r_i = 0$, and hence in the matrix notation as $\mathbf{A}^T \mathbf{r} = \mathbf{0}$.

Now,

$$\mathbf{0} = \mathbf{A}^T \mathbf{r} = \mathbf{A}^T (\mathbf{Ax} - \mathbf{b}) = \mathbf{A}^T \mathbf{Ax} - \mathbf{A}^T \mathbf{b}.$$

Thus the minimizing \mathbf{x} also solves the normal equations

$$\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b}.$$

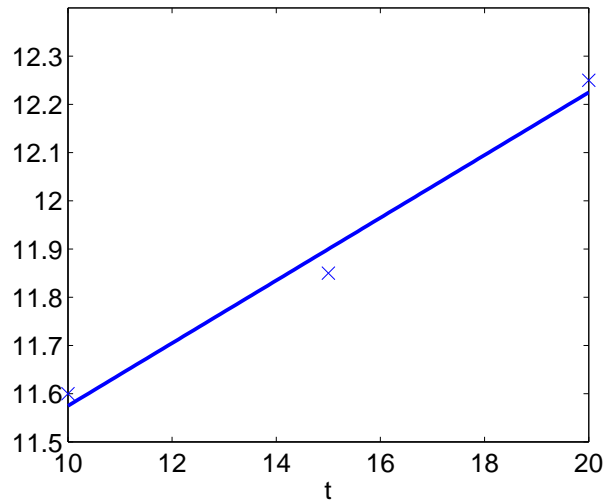


Figure 3: The data and the least-squares fit.

(b) The overdetermined system is $\mathbf{Ax} = \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & 10 \\ 1 & 15 \\ 1 & 20 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 11.60 \\ 11.85 \\ 12.25 \end{bmatrix}$$

The system is not consistent. The following solutions are obtained if we only use two equations at a time:

- $x_1 = 11.1$, $x_2 = 0.05$ if the third equation is ignored.
- $x_1 = 10.95$, $x_2 = 0.65$ if the second equation is ignored.
- $x_1 = 10.65$, $x_2 = 0.08$ if the first equation is ignored.

There is no particular reason to prefer any of these choices.

The normal equations are $\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b}$, i.e.,

$$\begin{bmatrix} 1 & 1 & 1 \\ 10 & 15 & 20 \end{bmatrix} \begin{bmatrix} 1 & 10 \\ 1 & 15 \\ 1 & 20 \end{bmatrix} \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 10 & 15 & 20 \end{bmatrix} \begin{bmatrix} 11.60 \\ 11.85 \\ 12.25 \end{bmatrix}.$$

The solution is $x_1 = 10.925$, $x_2 = 0.065$. The data and the fit are shown in Figure 3.