1.1 Problem Sheet - LinAlg 2-1 Projection and Least Squares

Given the data set of measured values $\underline{\mathbf{b}} = (0, 8, 8, 20)$, corresponding to $\underline{\mathbf{t}} = (0, 1, 3, 4)$, set up and solve the normal equations $A^T A\hat{\underline{\mathbf{x}}} = A^T \underline{\mathbf{b}}$.

1. Find the line of best fit $y = C + Dt$ for the above data set, with the line going through the points $(0, y_1), (1, y_2), (3, y_3)$ and $(4, y_4)$.

   Solution: The line has equations $y = C + Dt$ and we need C, D, which need to satisfy

$$\begin{aligned} 0 &= C + 0 \cdot D \\ 8 &= C + 1 \cdot D \\ 8 &= C + 3 \cdot D \\ 20 &= C + 4 \cdot D \end{aligned}$$

   and in matrix form we have $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$:

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 0 \\ 8 \\ 8 \\ 20 \end{pmatrix}.$$

   This is an overdetermined system with no solution, we seek the best approximation $\hat{C}, \hat{D}$. Begin with

$$A^T A = \begin{pmatrix} 4 & 8 \\ 8 & 26 \end{pmatrix} \Longrightarrow A^T \underline{\mathbf{b}} = \begin{pmatrix} 36 \\ 112 \end{pmatrix}$$

   We solve $A^T A\hat{\underline{\mathbf{x}}} = A^T \underline{\mathbf{b}}$ to give

$$\hat{\underline{\mathbf{x}}} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} \hat{C} \\ \hat{D} \end{pmatrix}$$

   so the line of best fit is $y = 1 + 4t$.

   Sketch the line of best fit along the data points indicating the errors $\underline{\mathbf{e}}$.

2. Obtain the errors $e_i = b_i - p_i$ and calculate the minimum $E = \sum e_i^2$

   Solution: Following the notation used in the notes, $\underline{\mathbf{p}} = A^T \underline{\mathbf{b}}$ are the points on the line of best fit, so

$$\underline{\mathbf{p}} = \begin{pmatrix} 1 \\ 5 \\ 13 \\ 17 \end{pmatrix} \Longrightarrow \underline{\mathbf{e}} = \underline{\mathbf{b}} - \underline{\mathbf{p}} = \begin{pmatrix} 0 \\ 8 \\ 8 \\ 20 \end{pmatrix} - \begin{pmatrix} 1 \\ 5 \\ 13 \\ 17 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ -5 \\ 3 \end{pmatrix}$$

   and $E = \sum e_i^2 = 44$. This is the minimal error for a line of best fit. We will see later how that is reduced.

3. Write down $E = \sum |A\underline{\mathbf{x}} - \underline{\mathbf{b}}|^2$ as a sum of four squares, the last one is $(C + 4D - 20)^2$. Obtain the partial derivatives w.r.t C and D, and set them equal to zero to find the minimum of E, thus obtaining the normal equations again.

   Solution: Begin with

$$E = \sum |A\underline{\mathbf{x}} - \underline{\mathbf{b}}|_i^2 = f(C, D) = (C + 0D - 0)^2 + (C + 1D - 8)^2 + (C + 3D - 8)^2 + (C + 4D - 20)^2$$

   Now differentiate and set partial derivatives equal to zero, to find the minimum of this function of two variables:

$$\frac{\partial E}{\partial C} = 2C + 2(C + D - 8) + 2(C` + 3D - 8) + 2(C + 4D - 20) = 0 \Rightarrow 4C + 8D = 36$$

and
$$\frac{\partial E}{\partial D} = 2(C+D-8)3 + 2(C` + 3D - 8)3 + 2(C+4D-20)4 = 0 \Rightarrow 8C + 26D = 112$$
or in matrix form
$$\begin{pmatrix} 4 & 8 \\ 8 & 26 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 36 \\ 112 \end{pmatrix}$$
which we recognize as
$$A^T A \hat{\underline{\mathbf{x}}} = A^T \hat{\underline{\mathbf{b}}}$$
which we had already solved giving $C, D = 1, 4$ as the values that minimize the sum of squared errors $E = f(C, D)$.

4. Find the closest line $y = Dt$ through the origin to the same four points. The matrix you need is now 4×1!. What do you expect the error E to be, compared to the previous case? Now confirm by calculating it!

Solution: Settign $C = 0$ we get a line through the origin $y = Dt$. The equations are now
$$\begin{aligned} 0 &= 0 \cdot D \\ 8 &= 1 \cdot D \\ 8 &= 3 \cdot D \\ 20 &= 4 \cdot D \end{aligned}$$
and in matrix form we have $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$:
$$\begin{pmatrix} 0 \\ 1 \\ 3 \\ 4 \end{pmatrix} \begin{pmatrix} D \end{pmatrix} = \begin{pmatrix} 0 \\ 8 \\ 8 \\ 20 \end{pmatrix}.$$
This time $A^T A = 26$ and matrices become superfluous. Likewise $A^T \underline{\mathbf{b}} = 112$ so the normal matrix equation $A^T A \hat{\underline{\mathbf{x}}} = A^T \underline{\mathbf{b}}$ is reduced to
$$26\hat{D} = 112$$
and the line of best fit through the origin is $y = \frac{56}{13}t$. We expect the error to be larger given that $y = 1 + 4t$ is *the* line of best fit.

Begin with the points on the line:
$$\underline{\mathbf{p}} = \begin{pmatrix} 0 \\ 56/13 \\ 168/13 \\ 224/13 \end{pmatrix} \Longrightarrow \underline{\mathbf{e}} = \underline{\mathbf{b}} - \underline{\mathbf{p}} = \frac{1}{13} \begin{pmatrix} -1 \\ 48 \\ -64 \\ 36 \end{pmatrix}$$
so that
$$E = \sum e_i^2 = \frac{1}{13^2}(1 + 48^2 + 64^2 + 36^2) \approx 45.5... > 44$$
as expected.

5. Find the best quadratic approximation $y = C + Dt + Ft^2$ to the same four points. Calculate the error $E = \sum e_i^2$ again. What do you notice?

Solution: Fitting a parabola $y = C + Dt + Ft^2$ through the points will reduce the error, as we can now get closer to the points than with a line. We set up the equations
$$\begin{aligned} 0 &= C + 0D + 0F \\ 8 &= C + 1D + F \\ 8 &= C + 3D + 9F \\ 20 &= C + 4D + 16F \end{aligned}$$

and in matrix form we have $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{pmatrix} \begin{pmatrix} C \\ D \\ F \end{pmatrix} = \begin{pmatrix} 0 \\ 8 \\ 8 \\ 20 \end{pmatrix}.$$

Hence the normal equation $A^T A\hat{\underline{\mathbf{x}}} = A^T\underline{\mathbf{b}}$ is

$$\begin{pmatrix} 4 & 8 & 26 \\ 8 & 26 & 92 \\ 26 & 92 & 338 \end{pmatrix} \begin{pmatrix} \hat{C} \\ \hat{D} \\ \hat{F} \end{pmatrix} = \begin{pmatrix} 36 \\ 112 \\ 400 \end{pmatrix}.$$

which we solve to find

$$\hat{\underline{\mathbf{x}}} = \begin{pmatrix} \hat{C} \\ \hat{D} \\ \hat{F} \end{pmatrix} = \begin{pmatrix} 2 \\ 4/3 \\ 2/3 \end{pmatrix}$$

and the parabola of best fit is $y = 2 + \frac{4}{3}t + \frac{2}{3}t^2$.

The points on the parabola give

$$\underline{\mathbf{p}} = \begin{pmatrix} 2 \\ 4 \\ 8 \\ 18 \end{pmatrix} \Longrightarrow \underline{\mathbf{e}} = \underline{\mathbf{b}} - \underline{\mathbf{p}} = \begin{pmatrix} 0 \\ 8 \\ 8 \\ 20 \end{pmatrix} - \begin{pmatrix} 2 \\ 4 \\ 8 \\ 18 \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \\ 0 \\ 2 \end{pmatrix}$$

so that $E = \sum e_i^2 = 24$, substantially less than for the line.

6. Find the best cubic approximation $y = C + Dt + Ft^2 + Gt^3$ to the same four points. Begin by setting up $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ and consider A before proceeding.

Solution: Fitting a cubic $y = C + Dt + Ft^2 + Gt^3$, we set up the equations

$$\begin{aligned} 0 &= C + 0D + 0F + 0G \\ 8 &= C + 1D + F + G \\ 8 &= C + 3D + 9F + 27G \\ 20 &= C + 4D + 16F + 64G \end{aligned}$$

and in matrix form we have $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{pmatrix} \begin{pmatrix} C \\ D \\ F \\ G \end{pmatrix} = \begin{pmatrix} 0 \\ 8 \\ 8 \\ 20 \end{pmatrix}.$$

The matrix is invertible, with solution

$$\begin{pmatrix} C \\ D \\ F \\ G \end{pmatrix} = \begin{pmatrix} 0 \\ 47/3 \\ -28/3 \\ 5/3 \end{pmatrix}$$

Using a cubic, we can go through the exact points, no approximation necessary. The principle is simple: we know that in the plane any two points lie on a unique line. This extends: any three non-collinear points lie on a unique parabola and if four points don't lie on a parabola, we can find a unique cubic through them.

1.2 Problem Sheet - LinAlg 2-2 - Gram-Schmidt and QR-decomposition

1. Given the orthogonal set of vectors $\underline{\mathbf{q}}_1, \underline{\mathbf{q}}_2, \ldots \underline{\mathbf{q}}_n$, show that they are linearly independent.

   Solution: To show linear independence, we require that the only solution to

   $$c_1\underline{\mathbf{q}}_1 + c_2\underline{\mathbf{q}}_2 + \ldots + c_n\underline{\mathbf{q}}_n = \underline{\mathbf{0}} \qquad (*)$$

   is $c_i = 0$ for all $i = 1 \ldots n$. This is easily shown. Take the scalar product of $(*)$ with $\underline{\mathbf{q}}_1$:

   $$\underline{\mathbf{q}}_1 \cdot (c_1\underline{\mathbf{q}}_1 + c_2\underline{\mathbf{q}}_2 + \ldots + c_n\underline{\mathbf{q}}_n) = \underline{\mathbf{q}}_1 \cdot \underline{\mathbf{0}} \Rightarrow c_1(\underline{\mathbf{q}}_1 \cdot \underline{\mathbf{q}}_1) + 0 + 0 + \ldots 0 = 0 \Rightarrow c_1 = 0$$

   given the orthogonal relation and $\underline{\mathbf{q}}_1 \cdot \underline{\mathbf{q}}_1 \neq 0$. Successively, taking the scalar product of $(*)$ with $\underline{\mathbf{q}}_2, \underline{\mathbf{q}}_3, \ldots$ yields $c_2 = 0$, $c_3 = 0$ and so on. All c_i are zero, so the vectors are linearly independent.

2. Consider the following basis for $\mathbf{R}^4$:

   $$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -2 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 18 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

   Apply the Gram-Schmidt process to this basis to produce an orthonormal basis for $\mathbf{R}^4$.

   Solution: Let

   $$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 2 \\ 2 \\ 1 \\ 0 \end{bmatrix}.$$

   Then,

   $$\begin{aligned} \mathbf{u}_2 &= \mathbf{v}_2 - \frac{\mathbf{u}_1 \cdot \mathbf{v}_2}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 \\ &= \begin{bmatrix} -2 \\ 1 \\ 2 \\ 0 \end{bmatrix} - \frac{0}{9}\begin{bmatrix} 2 \\ 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 2 \\ 0 \end{bmatrix} \end{aligned}$$

   (Note that $\mathbf{v}_1$ and $\mathbf{v}_2$ were orthogonal to begin with) and

   $$\begin{aligned} \mathbf{u}_3 &= \mathbf{v}_3 - \frac{\mathbf{u}_1 \cdot \mathbf{v}_3}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_3}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 \\ &= \begin{bmatrix} 18 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \left(\frac{36}{9}\right)\begin{bmatrix} 2 \\ 2 \\ 1 \\ 0 \end{bmatrix} - \left(\frac{-36}{9}\right)\begin{bmatrix} -2 \\ 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -4 \\ 4 \\ 0 \end{bmatrix}. \end{aligned}$$

   (Note that we can check that $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$, $\mathbf{u}_2 \cdot \mathbf{u}_3 = 0$, and $\mathbf{u}_1 \cdot \mathbf{u}_3 = 0$.) Finally,

   $$\begin{aligned} \mathbf{u}_4 &= \mathbf{v}_4 - \frac{\mathbf{u}_1 \cdot \mathbf{v}_4}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_4}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_4}{\mathbf{u}_3 \cdot \mathbf{u}_3}\mathbf{u}_3 \\ &= \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \left(\frac{5}{9}\right)\begin{bmatrix} 2 \\ 2 \\ 1 \\ 0 \end{bmatrix} - \left(\frac{1}{9}\right)\begin{bmatrix} -2 \\ 1 \\ 2 \\ 0 \end{bmatrix} - \frac{2}{36}\begin{bmatrix} 2 \\ -4 \\ 4 \\ 0 \end{bmatrix}. = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \end{aligned}$$

Note that this last result could have been deduced merely looking at the original four vectors: $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ all have zero entry in the x_4 direction, as do their linear combinations $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$: the vector $[0, 0, 0, 1]$ is guaranteed to be orthogonal to all of them. We now have $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3, \mathbf{u}_4$, an orthogonal basis for $\mathbf{R}^4$.

To get an orthonormal basis, we need to normalize each of $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ to have unit length; $\mathbf{u}_4$ is already a unit vector. So

$$\mathbf{q}_1 = \frac{1}{\|\mathbf{u}_1\|}\mathbf{u}_1 = \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{q}_2 = \frac{1}{\|\mathbf{u}_2\|}\mathbf{u}_2 = \frac{1}{3}\begin{bmatrix} -2 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{q}_3 = \frac{1}{\|\mathbf{u}_3\|}\mathbf{u}_3 = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{q}_4 = \mathbf{u}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

form the required orthonormal basis.

3. For the Gram-Schmidt process, given three independent vectors $\underline{\mathbf{a}}, \underline{\mathbf{b}}, \underline{\mathbf{c}}$ we construct an orthogonal set as follows:

$$\underline{\mathbf{x}}_1 = \underline{\mathbf{a}}\,, \quad \underline{\mathbf{x}}_2 = \underline{\mathbf{b}} - \left(\frac{\underline{\mathbf{x}}_1^T\underline{\mathbf{b}}}{\underline{\mathbf{x}}_1^T\underline{\mathbf{x}}_1}\right)\underline{\mathbf{x}}_1\,, \quad \text{and} \quad \underline{\mathbf{x}}_3 = \underline{\mathbf{c}} \quad - \quad \left(\frac{\underline{\mathbf{x}}_1^T\underline{\mathbf{c}}}{\underline{\mathbf{x}}_1^T\underline{\mathbf{x}}_1}\right)\underline{\mathbf{x}}_1 \quad - \quad \left(\frac{\underline{\mathbf{x}}_2^T\underline{\mathbf{c}}}{\underline{\mathbf{x}}_2^T\underline{\mathbf{x}}_2}\right)\underline{\mathbf{x}}_2\,.$$

Show that $\underline{\mathbf{x}}_3$ is orthogonal to $\underline{\mathbf{x}}_1$ and $\underline{\mathbf{x}}_2$.

Solution: By direct substitution.

4. Obtain the $QR-$decomposition of

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

Solution: Using Gram-Schmidt on the columns of A, we obtain the orthonormal set

$$\frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix}, \quad \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \quad \text{and} \quad \frac{1}{2}\begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix},$$

giving the orthogonal matix

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

The matrix R by construction is

$$R = \begin{pmatrix} \underline{\mathbf{q}}_1^T\underline{\mathbf{a}}_1 & \underline{\mathbf{q}}_1^T\underline{\mathbf{a}}_2 & \underline{\mathbf{q}}_1^T\underline{\mathbf{a}}_3 & \underline{\mathbf{q}}_1^T\underline{\mathbf{a}}_4 \\ 0 & \underline{\mathbf{q}}_2^T\underline{\mathbf{a}}_2 & \underline{\mathbf{q}}_2^T\underline{\mathbf{a}}_3 & \underline{\mathbf{q}}_2^T\underline{\mathbf{a}}_4 \\ 0 & 0 & \underline{\mathbf{q}}_3^T\underline{\mathbf{a}}_3 & \underline{\mathbf{q}}_3^T\underline{\mathbf{a}}_4 \\ 0 & 0 & 0 & \underline{\mathbf{q}}_4^T\underline{\mathbf{a}}_4 \end{pmatrix}$$

where $\underline{\mathbf{q}}_i^T$ are the columns of Q and $\underline{\mathbf{a}}_i$ the columns of A. Computing the desired scalar products we obtain

$$R = \begin{pmatrix} 2 & 3 & 2 & \frac{11}{2} \\ 0 & 5 & -2 & \frac{3}{2} \\ 0 & 0 & 4 & \frac{5}{2} \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}$$

Multiply out QR to confirm.

5. For extra practice, take the four vectors in Q2 as the columns of a matrix A. You already have Q, now find R for the $QR-$decomposition of A.

1.3 Problem Sheet - LinAlg 2-3 - Eigenvalues, Eigenvectors and Diagonalization

1. Find the eigenvalues and eigenvectors of the following matrices:

(a) $A = \begin{pmatrix} -1 & 1 \\ 4 & 2 \end{pmatrix}$ (b) $B = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$ (c) $C = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix}$

Solution:

(a) Begin with eigenvalues, find the characteristic equation:

det$(A - \lambda I) = \begin{vmatrix} -1-\lambda & 1 \\ 4 & 2-\lambda \end{vmatrix} = \lambda^2 - \lambda - 6 = 0 \Rightarrow \lambda_1 = 3\,, \lambda_2 = -2\,.$

To find the eigenvectors, solve $(A - \lambda I)\underline{\mathbf{x}} = \underline{\mathbf{0}}$ for each eigenvalue.

For $\lambda_1 = 3: \begin{pmatrix} -4 & 1 \\ 4 & -1 \end{pmatrix} \underline{\mathbf{x}} = \underline{\mathbf{0}} \Rightarrow -4x + y = 0$

There is an expected free variable, choose $x = 1 \Rightarrow y = 4$ and the eigenvector is $\underline{\mathbf{x}}_1 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$, or any non-zero multiple of this.

For $\lambda_2 = -2: \begin{pmatrix} 1 & 1 \\ 4 & 4 \end{pmatrix} \underline{\mathbf{x}} = \underline{\mathbf{0}} \Rightarrow x + y = 0$

There is an expected free variable, choose $x = 1 \Rightarrow y = -1$ and the eigenvector is $\underline{\mathbf{x}}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, or any non-zero multiple of this.

(b) Again, det$(A - \lambda I) = \begin{vmatrix} 1-\lambda & -2 \\ 2 & 1-\lambda \end{vmatrix} = (1-\lambda^2) + 4 = 0 \Rightarrow \lambda_1 = 1 + 2i\,, \lambda_2 = 1 - 2i\,,$ and with complex eigenvalues, we expect complex eigenvectors, but again solve $(B - \lambda I)\underline{\mathbf{x}} = \underline{\mathbf{0}}$ for each eigenvalue.

For $\lambda_1 = 1 + 2i: (B - \lambda_1 I)\underline{\mathbf{x}} = \underline{\mathbf{0}} \Rightarrow \begin{pmatrix} 2i & -2 \\ 2 & 2i \end{pmatrix} \underline{\mathbf{x}} = \underline{\mathbf{0}} \Rightarrow 2ix - 2y = 0$ or $2x + 2iy = 0$, both give $x = iy$. Choose $y = 1 \Rightarrow x = i$ and the eigenvector is $\underline{\mathbf{x}}_1 = \begin{pmatrix} i \\ 1 \end{pmatrix}$, or any non-zero multiple of this.

For $\lambda_2 = 1 - 2i: \begin{pmatrix} -2i & -2 \\ 2 & -2i \end{pmatrix} \underline{\mathbf{x}} = \underline{\mathbf{0}} \Rightarrow -2ix - 2y = 0$ or $2x - 2iy = 0$, both give $x = -iy$. Choose $y = 1 \Rightarrow x = -i$ and the eigenvector is $\underline{\mathbf{x}}_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$, or any non-zero multiple of this.

(c) For the 2×2 case, finding the eigenvalues is easy, but for higher it's useful to try row/column operations to try to find a factorized characteristic polynomial: det$(C - \lambda I)$

$$= \begin{vmatrix} 1-\lambda & 0 & -1 \\ 1 & 2-\lambda & 1 \\ 2 & 2 & 3-\lambda \end{vmatrix} \underset{R_1+R_2}{=} \begin{vmatrix} 2-\lambda & 2-\lambda & 0 \\ 1 & 2-\lambda & 1 \\ 2 & 2 & 3-\lambda \end{vmatrix} \underset{C_2-C_1}{=} \begin{vmatrix} 2-\lambda & 0 & 0 \\ 1 & 1-\lambda & 1 \\ 2 & 0 & 3-\lambda \end{vmatrix}$$

$$= (2-\lambda) \begin{vmatrix} 1-\lambda & 1 \\ 0 & 3-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda)(3-\lambda) = 0 \Rightarrow \lambda = 1, 2, 3\,.$$

To find the eigenvectors, we solve $(C - \lambda I)\underline{\mathbf{x}} = \underline{\mathbf{0}}$ for each eigenvalue, but may need row operations. For $\lambda_1 = 1$:

$\begin{pmatrix} 0 & 0 & -1 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix} \underline{\mathbf{x}} = \underline{\mathbf{0}}$. Here row 1 gives $z = 0$ and so no row ops needed to see that $x + y + z = x + y = 0$ for the other two rows. Choose $x = 1 \Rightarrow y = -1$ and the eigenvector is $\underline{\mathbf{x}}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, or any non-zero multiple. For $\lambda_2 = 2:$

$\begin{pmatrix} -1 & 0 & -1 \\ 1 & 0 & 1 \\ 2 & 2 & 1 \end{pmatrix} \underline{\mathbf{x}} = \underline{\mathbf{0}}$. Again, no row operations are needed as both row one and two give $x + z = 0$, giving a free variable. Choose $z = 1 \Rightarrow x = -1$, then row three gives $2x + 2y + z = -2 + 2y + 1 = 0 \Rightarrow y = 1/2$ and the eigenvector is $\underline{\mathbf{x}}_2 = \begin{pmatrix} -1 \\ \frac{1}{2} \\ 1 \end{pmatrix}$, or any non-zero multiple. For $\lambda_3 = 3$:

$$\begin{pmatrix} -2 & 0 & -1 \\ 1 & -1 & 1 \\ 2 & 2 & 0 \end{pmatrix} \underset{R_1 \leftrightarrow R_2}{\sim} \begin{pmatrix} 1 & -1 & 1 \\ -2 & 0 & -1 \\ 2 & 2 & 0 \end{pmatrix} \underset{\substack{R_2+2R_1 \\ R_3-2R_1}}{\sim} \begin{pmatrix} 1 & -1 & 1 \\ 0 & -2 & 1 \\ 0 & 4 & -2 \end{pmatrix} \underline{\mathbf{x}} = \underline{\mathbf{0}}\,.$$

There is no need for the augmented coefficient matrix here, as $(A : \underline{\mathbf{b}}) = (A : \underline{\mathbf{0}})$ in every case, so row ops don't affect the fourth column. We can stop here as both rows two and three give $2y - z = 0$, and we choose the free variable $y = 1 \Rightarrow z = 2$. Row one then gives $x - y + z = x - 1 + 2 = 0 \Rightarrow x = -1$ and the eigenvector is $\underline{\mathbf{x}}_3 = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$, or any non-zero multiple.

2. Obtain the orthogonal diagonalization of $A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$.

Solution: We begin by finding eigenvalues and eigenvectors. First solve $0 =$det$(A - \lambda I) =$

$$= \begin{vmatrix} 1-\lambda & 2 & 2 \\ 2 & 1-\lambda & 2 \\ 2 & 2 & 1-\lambda \end{vmatrix} \underset{C_1 - C_2}{=} \begin{vmatrix} -1-\lambda & 2 & 2 \\ 1+\lambda & 1-\lambda & 2 \\ 0 & 2 & 1-\lambda \end{vmatrix} \underset{R_2 + R_1}{=} \begin{vmatrix} -1-\lambda & 2 & 2 \\ 0 & 3-\lambda & 4 \\ 0 & 2 & 1-\lambda \end{vmatrix}$$

$$= (-1-\lambda)[((3-\lambda)(1-\lambda) - 8] = -(1+\lambda)(\lambda^2 - 4\lambda - 5) = -(\lambda+1)^2(\lambda - 5)\,,$$

so the eigenvalues are $\lambda_1 = 5$ (simple) and $\lambda_2 = -1$ (repeated).

The eigenvector for λ_1 is found as before, solving $(A - 5I)\underline{\mathbf{x}} = \underline{\mathbf{0}}$:

$\begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix} \underline{\mathbf{x}} = \underline{\mathbf{0}}$, Gauss $\Rightarrow x = y = z$, so $\underline{\mathbf{x}}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, normalize $\Rightarrow \hat{\underline{\mathbf{x}}}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

The eigenvectors for λ_2 are found as usual solving $(A + I)\underline{\mathbf{x}} = \underline{\mathbf{0}}$: $\begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \underline{\mathbf{x}} = \underline{\mathbf{0}}\,, \Rightarrow x + y + z = 0\,.$

There are two free variables as expected, giving two linearly independent eigenvectors. We can guarantee linear independence by choosing $y = 1, z = 0 \Rightarrow x = -1$ for one, and choosing $y = 1, z = 0 \Rightarrow x = -1$ for the other. So the two eigenvectors are $\underline{\mathbf{x}}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, and $\underline{\mathbf{x}}_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

As expected for a symmetric matrix, eigenvectors corresponding to distinct eigenvalues are orthogonal: $\underline{\mathbf{x}}_1 \cdot \underline{\mathbf{x}}_2 = 0$ and $\underline{\mathbf{x}}_1 \cdot \underline{\mathbf{x}}_3 = 0$, but $\underline{\mathbf{x}}_2 \cdot \underline{\mathbf{x}}_3 \neq 0$, as these correspond to the same eigenvalue.

We use the technique of projection to obtain two orthogonal eigenvectors from x_2 and x_3. Recall that if $\mu\underline{\mathbf{x}}_3$ is the projection of $\underline{\mathbf{x}}_2$ onto $\underline{\mathbf{x}}_3$, then $\underline{\mathbf{x}}_2 - \mu\underline{\mathbf{x}}_3$ is orthogonal to $\underline{\mathbf{x}}_3$. Obtain

$$\mu = \frac{\underline{\mathbf{x}}_2 \cdot \underline{\mathbf{x}}_3}{\underline{\mathbf{x}}_3 \cdot \underline{\mathbf{x}}_3} = \frac{1}{2} \Rightarrow \underline{\mathbf{x}}_4 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

Check that (i) $\underline{\mathbf{x}}_4$ is an eigenvector for A, and (ii) $\underline{\mathbf{x}}_4 \cdot \underline{\mathbf{x}}_1 = 0$. Finally, normalize and

$\hat{\underline{\mathbf{x}}}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\quad \hat{\underline{\mathbf{x}}}_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, $\quad \hat{\underline{\mathbf{x}}}_4 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ is an orthonormal set of eigenvectors of A.

Note that the choice of projection was free. We could hace chosen the projection of $\underline{\mathbf{x}}_3$ onto $\underline{\mathbf{x}}_2$ and obtained orthogonal eigenvectors $\underline{\mathbf{x}}_2$ and $\underline{\mathbf{x}}_3 - \mu\underline{\mathbf{x}}_2$.

For the orthogonal diagonalization, we find an orthogonal matrix P such that $A = PDP^T$ where D is a diagonal matrix whose diagonal elements are the eigenvalues of A, in the same order as the orthonormal eigenvectors are the columns of P, so that

$$P = \frac{1}{\sqrt{6}}\begin{pmatrix} \sqrt{2} & -\sqrt{3} & 1 \\ \sqrt{2} & 0 & -2 \\ \sqrt{2} & \sqrt{3} & 1 \end{pmatrix}, \quad \text{and} \quad D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

3. (a) A system of coupled first order ODEs is given by

$$\dot{x} = -\frac{2}{3}x + \frac{1}{3}y - 4\,, \quad \dot{y} = \frac{1}{3}x - \frac{2}{3}y + 2$$

where $x(0) = y(0) = 1$.

Write the system in matrix form and use diagonalisation to decouple the equations and find $x(t)$ and $y(t)$.

Solution: In matrix form we have

$$\dot{\underline{\mathbf{x}}} = A\underline{\mathbf{x}} + \underline{\mathbf{u}} \Rightarrow \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -4 \\ 2 \end{pmatrix}$$

The eigenvalues are $-\frac{1}{3}, -1$ with corresponding normalized eigenvectors $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ so the diagonalization is $A = S\Lambda S^{-1}$ with diagonalizing/diagonal matrices

$$S = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \Lambda = \begin{pmatrix} -\frac{1}{3} & 0 \\ 0 & -1 \end{pmatrix}$$

We now define a new variable $\underline{\mathbf{z}}$ satisfying $\underline{\mathbf{z}} = S^{-1}\underline{\mathbf{x}}$ so that

$$\dot{\underline{\mathbf{z}}} = S^{-1}\dot{\underline{\mathbf{x}}} = S^{-1}A\underline{\mathbf{x}} + S^{-1}\underline{\mathbf{u}} = S^{-1}ASS^{-1}\underline{\mathbf{x}} + S^{-1}\underline{\mathbf{u}} = \Lambda S^{-1}\underline{\mathbf{x}} + S^{-1}\underline{\mathbf{u}} = \Lambda\underline{\mathbf{z}} + S^{-1}\underline{\mathbf{u}}$$

As A is symmetric, $S^T = S^{-1}$ and we can write

$$\dot{\underline{\mathbf{z}}} = \begin{pmatrix} \dot{z}_1 \\ \dot{z}_2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} -4 \\ 2 \end{pmatrix}$$

giving the decoupled first order linear ODEs

$$\dot{z}_1 = -\frac{1}{3}z_1 - \sqrt{2} \quad \text{and} \quad \dot{z}_2 = -z_2 - 3\sqrt{2}$$

with solutions

$$z_1 = c_1 e^{-\frac{1}{3}t} - 3\sqrt{2}\,, \quad z_2 = c_2 e^{-t} - 3\sqrt{2}$$

Now

$$\underline{\mathbf{x}} = S\underline{\mathbf{z}} \Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} c_1 e^{-\frac{1}{3}t} - 3\sqrt{2} \\ c_2 e^{-t} - 3\sqrt{2} \end{pmatrix}$$

Now implement the initial conditions to get $c_1 = 4\sqrt{2}$ and $c_2 = 3\sqrt{2}$ so the solution is

$$x(t) = -6 + 3e^{-t} + 4e^{-\frac{1}{3}t} \qquad y(t) = -3e^{-t} + 4e^{-\frac{1}{3}t}$$

(b) A system of coupled first order ODEs is given by

$$\dot{x} + 5x + 2y = e^{-t}, \quad \dot{y} + 2x + 2y = 0, \quad x(0) = 1, \;\; y(0) = 0.$$

Write the system in matrix form and use diagonalisation to decouple the system and find $x(t)$ and $y(t)$.

Solution: DiY!

$$x(t) = \frac{1}{5}\left(\frac{9}{5} + t\right)e^{-t} + \frac{16}{25}e^{-6t} \qquad y(t) = -\frac{2}{5}\left(\frac{4}{5} + t\right)e^{-t} + \frac{8}{25}e^{-6t}$$

4. (from Riley, Hobson and Bence, Ex 8.19) Given that A is a real symmetric matrix with normalised (unit) eigenvectors $\underline{\mathbf{e}}_i$, the eigenvectors can form a basis and any vector can be written as a linear combination of these:

$$\underline{\mathbf{x}} = \sum_i \alpha_i \underline{\mathbf{e}}_i \,.$$

If the vector $\underline{\mathbf{x}}$ is the solution of

$$A\underline{\mathbf{x}} - \mu\underline{\mathbf{x}} = \underline{\mathbf{v}} \,, \qquad (*)$$

obtain the coefficients α_i, involved in the expansion.

Here μ is a given constant and $\underline{\mathbf{v}}$ is a given vector. You may find it useful to recall the derivation of the Euler formulae for real Fourier coefficients.

(a) Solve $(*)$ when $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$, $\mu = 2$ and $\underline{\mathbf{v}} = (1,2,3)^T$.

(b) With the same A as in (a), and $\mu = 1$, would $(*)$ have a solution if (i) $\underline{\mathbf{v}} = (1,2,3)^T$, (ii) $\underline{\mathbf{v}} = (2,2,3)^T$?

Solution: Begin by substituting $\underline{\mathbf{x}} = \sum_i \alpha_i \underline{\mathbf{e}}_i$ into $(*)$:

$$A\sum_i \alpha_i \underline{\mathbf{e}}_i - \mu \sum_i \alpha_i \underline{\mathbf{e}}_i = \underline{\mathbf{v}} \Longrightarrow \sum_i \alpha_i (A\underline{\mathbf{e}}_i) - \sum_i \mu\alpha_i \underline{\mathbf{e}}_i = \underline{\mathbf{v}} \Longrightarrow \sum_i (\lambda_i - \mu)\alpha_i \underline{\mathbf{e}}_i = \underline{\mathbf{v}} \,, \quad (**)$$

where λ_i is the eigenvalue corresponding to $\underline{\mathbf{e}}_i$. Now take the scalar product of both sides with $\underline{\mathbf{e}}_j$:

$$\underline{\mathbf{e}}_j \cdot \sum_i (\lambda_i - \mu)\alpha_i \underline{\mathbf{e}}_i = \sum_i (\lambda_i - \mu)\alpha_i \underline{\mathbf{e}}_j \cdot \underline{\mathbf{e}}_i = \underline{\mathbf{e}}_j \cdot \underline{\mathbf{v}} \,,$$

And as the eigenvectors form an orthonormal set, we have $\underline{\mathbf{e}}_j \cdot \underline{\mathbf{e}}_i = 1$ when $i = j$ and zero for all other i. The sum vanishes but for one term and we have

$$(\lambda_j - \mu)\alpha_j = \underline{\mathbf{e}}_j \cdot \underline{\mathbf{v}} \Longrightarrow \alpha_j = \frac{\underline{\mathbf{e}}_j \cdot \underline{\mathbf{v}}}{\lambda_j - \mu} \,.$$

Note that if $\mu = \lambda_j$ for any j, then $(**)$ becomes

$$\sum_{i \neq j} (\lambda_i - \mu)\alpha_i \underline{\mathbf{e}}_i = \underline{\mathbf{v}} \,,$$

so for a solution to exist, $\underline{\mathbf{v}}$ must be a linear combination of eigenvectors excluding $\underline{\mathbf{e}}_j \Longrightarrow \underline{\mathbf{e}}_j \cdot \underline{\mathbf{v}} = 0$. So unless $\underline{\mathbf{e}}_j \cdot \underline{\mathbf{v}} = 0$, there will be no solution when $\mu = \lambda_j$.

(a) The eigenvalues of the given matrix are obtained by taking

det$(A - \lambda I) = 0 \Longrightarrow (3 - \lambda)[(2-\lambda)^2 - 1] = 0 \Longrightarrow \lambda = 1$ (simple) and $\lambda = 2$ (double).

A possible set of orthonormal eigenvectors is

For $\lambda = 3\,, \underline{\mathbf{e}}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ and $\underline{\mathbf{e}}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$. For $\lambda = 1\,, \underline{\mathbf{e}}_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$.

With $\mu = 2$ and $\underline{\mathbf{v}} = (1,2,3)^T$ we first find the coefficients

$$\alpha_1 = \frac{(0,0,1)\cdot(1,2,3)}{3-2} = 3\,, \quad \alpha_2 = \frac{\frac{1}{\sqrt{2}}(1,1,0)\cdot(1,2,3)}{3-2} = \frac{3}{\sqrt{2}}\,, \quad \alpha_3 = \frac{\frac{1}{\sqrt{2}}(1,-1,0)\cdot(1,2,3)}{1-2} = \frac{1}{\sqrt{2}}\,.$$

Then the solution vector is

$$\underline{\mathbf{x}} = 3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \frac{3}{\sqrt{2}}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}\,.$$

(b) If $\mu = 1$, then it is equal to the third eigenvalue and a solution is only possible if $\underline{\mathbf{e}}_3 \cdot \underline{\mathbf{v}} = 0$.
For (i) $\underline{\mathbf{v}} = (1,2,3)^T$, $\underline{\mathbf{e}}_3 \cdot \underline{\mathbf{v}} = -1/\sqrt{2}$, so no solution is possible.
For (ii) $\underline{\mathbf{v}} = (2,2,3)^T$, $\underline{\mathbf{e}}_3 \cdot \underline{\mathbf{v}} = 0$, so a solution is possible. We now have

$$\alpha_1 = \frac{(0,0,1)\cdot(2,2,3)}{3-1} = \frac{3}{2}\,, \quad \alpha_2 = \frac{\frac{1}{\sqrt{2}}(1,1,0)\cdot(2,2,3)}{3-1} = \sqrt{2}\,,$$

and the solution is

$$\underline{\mathbf{x}} = \frac{3}{2}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \sqrt{2}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \frac{3}{2} \end{pmatrix}\,.$$

1.4 Problem Sheet - LinAlg 2-4 - Symmetric Matrices and SVD

1. Find the SVD for $A = \begin{pmatrix} 2 & 2 & 0 \\ -1 & 1 & 0 \end{pmatrix}$, using the alternative, and longer, $A^T A$. This was solved in lectures using AA^T.

Solution: Begin by multiplying to obtain

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

a singular matrix, which we expect. There are only two non-zero eigenvalues, as seen in the AA^T approach.

Now find the eigenvalues:

$$\det(A^T A - \lambda I) = \begin{vmatrix} 5-\lambda & 3 & 0 \\ 3 & 5-\lambda & 0 \\ 0 & 0 & -\lambda \end{vmatrix} = -\lambda[(5-\lambda)^2 - 9] = -\lambda(\lambda^2 - 10\lambda + 16) = 0$$

giving $\lambda = 0, 2, 8$, as expected from the previous solution.

Eigenvectors are found next, giving the columns of V. For $\lambda_1 = 8$:

$$(A^T A - 8I)\underline{\mathbf{x}} = \begin{pmatrix} -3 & 3 & 0 \\ 3 & -3 & 0 \\ 0 & 0 & -8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \underline{\mathbf{0}} \Rightarrow z = 0, x = y$$

so the normalized eigenvector is $\underline{\mathbf{v}}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$. Similarly, for $\lambda_2 = 2$:

$$(A^T A - 2I)\underline{\mathbf{x}} = \begin{pmatrix} 3 & 3 & 0 \\ 3 & 3 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \underline{\mathbf{0}} \Rightarrow z = 0, x = -y$$

so the normalized eigenvector is $\underline{\mathbf{v}}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$. Finally, for $\lambda_3 = 0$:

$$(A^T A - 0I)\underline{\mathbf{x}} = \begin{pmatrix} 5 & 3 & 0 \\ 3 & 5 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \underline{\mathbf{0}} \Rightarrow 5x + 3y = 0\ , \quad 3x + 5y = 0$$

with solution $x = y = 0$ and z free, so we choose $z = 1$, giving the normalized eigenvector $\underline{\mathbf{v}}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Note that all three eigenvectors are the same as previously obtained. The first two, corresponding to non-zero eigenvalues, are a basis for the row space, $C(A^T)$, while $\underline{\mathbf{v}}_3$, corresponding to the zero eigenvalue, is a basis for the 1-D Nullspace, $N(A)$. Taken together they form

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

We can now find the vectors $\underline{\mathbf{u}}_i$ directly, using the relation

$$\underline{\mathbf{u}}_i = \frac{1}{\sigma_i} A \underline{\mathbf{v}}_i\ ,$$

seen in lectures/notes. The singular values are the roots of the non-zero eigenvalues:

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{8} \quad \text{and} \quad \sigma_2 = \sqrt{\lambda_2} = \sqrt{2}$$

so that

$$\underline{\mathbf{u}}_1 = \frac{1}{\sigma_1} A \underline{\mathbf{v}}_1 \;=\; \frac{1}{\sqrt{8}} \begin{pmatrix} 2 & 2 & 0 \\ -1 & 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \;=\; \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and similarly,

$$\underline{\mathbf{u}}_2 = \frac{1}{\sigma_2} A \underline{\mathbf{v}}_2 \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & 2 & 0 \\ -1 & 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \;=\; \begin{pmatrix} 0 \\ 1 \end{pmatrix} ,$$

giving a basis for the column space, $C(A)$. There are no further $\underline{\mathbf{u}}_i$, so we see that the left Nullspace $N(A^T)$ consists of the zero-vector. They give U as the 2×2 identity matrix, and the SVD is complete:

$$A = \begin{pmatrix} 2 & 2 & 0 \\ -1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{8} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \;=\; U \Sigma V^T$$

2. Given $A = \begin{pmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{pmatrix}$, find (a) the Singular Value Decomposition, and (b) the pseudoinverse.

Solution: (a) As A is 3×2, for the faster solution, we need $A^T A$ which is 2×2:

$$A^T A = \begin{pmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 11 & 1 \\ 1 & 11 \end{pmatrix}$$

Find the eigenvalues:

$$\det(A^T A - \lambda I) = \begin{vmatrix} 11 - \lambda & 1 \\ 1 & 11 - \lambda \end{vmatrix} = (11 - \lambda)^2 - 1 = \lambda^2 - 22\lambda + 120 = 0 \Rightarrow \lambda = 10, 12 \; .$$

The eigenvectors of $A^T A$ are $\underline{\mathbf{v}}_i$, a basis for the row space, $C(A^T)$. For $\lambda_1 = 12$:

$$(A^T A - 12I)\underline{\mathbf{x}} \;=\; \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \underline{\mathbf{0}} \;\Rightarrow x = y$$

and the normalized eigenvector is $\underline{\mathbf{v}}_1 \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Similarly, for $\lambda_2 = 10$:

$$(A^T A - 10I)\underline{\mathbf{x}} \;=\; \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \underline{\mathbf{0}} \;\Rightarrow x = -y$$

and the normalized eigenvector is $\underline{\mathbf{v}}_2 \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. And we have

$$V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} .$$

Next, we can use the relation

$$\underline{\mathbf{u}}_i = \frac{1}{\sigma_i} A \underline{\mathbf{v}}_i \; ,$$

to find $\underline{\mathbf{u}}_i$ for $i = 1, 2$, while $\underline{\mathbf{u}}_3$ will be a solution of $A^T\underline{\mathbf{u}}_3 = \underline{\mathbf{0}}$, forming a basis for the left Nullspace. We have the singular values:

$$\sigma_1 = \sqrt{\lambda_1} = 2\sqrt{3} \quad \text{and} \quad \sigma_2 = \sqrt{\lambda_2} = \sqrt{10}$$

so that

$$\underline{\mathbf{u}}_1 = \frac{1}{\sigma_1}A\underline{\mathbf{v}}_1 \;=\; \frac{1}{2\sqrt{3}}\begin{pmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \;=\; \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix},$$

a unit vector, as expected. Similarly,

$$\underline{\mathbf{u}}_2 = \frac{1}{\sigma_2}A\underline{\mathbf{v}}_2 \;=\; \frac{1}{\sqrt{10}}\begin{pmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix} \;=\; \frac{1}{\sqrt{5}}\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix},$$

also a unit vector. Together they form a basis for the column space $C(A)$. We find $\underline{\mathbf{u}}_3$, a basis for the left nullspace, solving

$$A^T\underline{\mathbf{u}}_3 = \begin{pmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \underline{\mathbf{0}} \Rightarrow \underline{\mathbf{u}}_3 = \frac{1}{\sqrt{30}}\begin{pmatrix} -1 \\ -2 \\ 5 \end{pmatrix},$$

(detail: DIY) so that

$$U = \begin{pmatrix} \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{5}{\sqrt{30}} \end{pmatrix}$$

and the SVD is

$$A = \begin{pmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{5}{\sqrt{30}} \end{pmatrix}\begin{pmatrix} 2\sqrt{3} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \;=\; U\Sigma V^T.$$

For further practice, obtain the 3×3 matrix AA^T, find its eigenvalues, two of which will be the same as those of A^TA, and one extra eigenvalue, which will be zero. The eigenvectors will be the same vectors $\underline{\mathbf{u}}_i$ found above. Then finish by obtaining $\underline{\mathbf{v}}_i = \dfrac{1}{\sigma_i}A^T\underline{\mathbf{u}}_i$, for the non-zero eigenvalues, obtaining the same $\underline{\mathbf{v}}_i$.

(b) For an $m \times n$ matrix A, with SVD given by $A = U\Sigma V^T$, the pseudo-inverse is the $n \times m$ matrix $A^+ = V\Sigma^+U^T$, where the $n \times m$ matrix Σ^+ has the reciprocals of the singular values $\sigma_i$on the main diagonal, in the same order as in Σ. So we have already done all the work:

$$A^+ \;=\; V\Sigma^+U^T \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{2\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{10}} & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ -\frac{1}{\sqrt{30}} & -\frac{2}{\sqrt{30}} & \frac{5}{\sqrt{30}} \end{pmatrix} = \frac{1}{60}\begin{pmatrix} 17 & 4 & 5 \\ -7 & 16 & 5 \end{pmatrix}.$$

3. Obtain the SVD for the matrix

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

Solution: It's 3×4, so for the faster route, we obtain

$$AA^T = \begin{pmatrix} 1 & 0 & 2 & 1 \\ -2 & 1 & -1 & 0 \\ -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 6 & -4 & 2 \\ -4 & 6 & 2 \\ 2 & 2 & 4 \end{pmatrix}$$

Next, find the eigenvalues, expanding the determinant by the first row.

$$\det(AA^T - \lambda I) = \begin{vmatrix} 6-\lambda & -4 & 2 \\ -4 & 6-\lambda & 2 \\ 2 & 2 & 4-\lambda \end{vmatrix} = (6-\lambda)[(6-\lambda)(4-\lambda)-4] + 4[-4(4-\lambda)-4] + 2[-8-2(6-\lambda)] = 0$$

Multiplying out, we obtain the characteristic polynomial $-\lambda^3 + 16\lambda^2 - 60\lambda = 0 \Rightarrow \lambda = 10, 6, 0$. We note that we have a zero eigenvalue: AA^T is positive semi-definite and singular. Looking back at A, we note that the rows are dependent: $R_1 + R_2 = R_3$, and A has rank 2, so two non-zero eigenvalues is consistent. The normalized eigenvectors (detail: DYI) of AA^T are

$$\lambda_1 = 10 \Rightarrow \underline{\mathbf{u}}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad \text{and} \quad \lambda_2 = 6 \Rightarrow \underline{\mathbf{u}}_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \quad \text{and} \quad \lambda_3 = 0 \Rightarrow \underline{\mathbf{u}}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

There's something new here! $\underline{\mathbf{u}}_1$ and $\underline{\mathbf{u}}_2$ are a basis for the column space $C(A)$ and $\underline{\mathbf{u}}_3$ is a basis for the left nullspace $N(A^T)$. Now, when we find $\underline{\mathbf{v}}_i$, we expect $\underline{\mathbf{v}}_1$ and $\underline{\mathbf{v}}_2$, corresponding to the non-zero eigenvalues to form a basis for the row space $C(A^T)$, and then *two* further vectors $\underline{\mathbf{v}}_3$ and $\underline{\mathbf{v}}_4$, a basis for the nullspace $N(A)$. In all the previously seen examples, because A had either full column or full row rank, the smaller of the two symmetric matrices AA^T or A^TA would be invertible and have only non-zero eigenvalues. As a result, the first set of vectors we found would be a basis for the row or column space, which would be the whole of $\mathbb{R}^m$ or $\mathbb{R}^n$, and the accompanying nullspace would have only the zero vector. This time, both nullspaces have more than the zero vector.

We obtain

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

Now, using $\underline{\mathbf{v}}_i = \dfrac{1}{\sigma_i} A^T \underline{\mathbf{u}}_i$, we get the first two of the $\underline{\mathbf{v}}_i$, for the singular values

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{10} \quad \text{and} \quad \sigma_2 = \sqrt{\lambda_2} = \sqrt{6}$$

so that

$$\underline{\mathbf{v}}_1 = \frac{1}{\sigma_1} A^T \underline{\mathbf{u}}_1 = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \frac{1}{2\sqrt{5}} \begin{pmatrix} 3 \\ -1 \\ 3 \\ 1 \end{pmatrix}$$

and

$$\underline{\mathbf{v}}_2 = \frac{1}{\sigma_2} A^T \underline{\mathbf{u}}_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Finally, $\underline{\mathbf{v}}_3$ and $\underline{\mathbf{v}}_4$, form a basis for the nullspace $N(A)$, so we solve $A\underline{\mathbf{x}} = \underline{\mathbf{0}}$:

$$\begin{pmatrix} 1 & 0 & 2 & 1 \\ -2 & 1 & -1 & 0 \\ -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \underline{\mathbf{0}} \,.$$

Using Gaussian elimination (DIY) we obtain

$$\underline{\mathbf{v}}_3^* = \begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \underline{\mathbf{v}}_4^* = \begin{pmatrix} -2 \\ -3 \\ 1 \\ 0 \end{pmatrix}$$

But we require V to be an orthogonal matrix, so the vectors need to form an orthogonal set. The first two, $\underline{\mathbf{v}}_{1,2}$ found using the singular values and $\underline{\mathbf{u}}_{1,2}$, are orthogonal because they correspond to distinct non-zero eigenvalues of A^TA. On the other hand $\underline{\mathbf{v}}_{3,4}$, forming the basis of the nullspace, correspond to the repeated eigenvalue $\lambda = 0$. In other words, we can obtain orthogonal eigenvectors, but we need to use projection, (or Gram-Schmidt, if there are more than two) on the first version. So, once more, we take $\underline{\mathbf{v}}_{3,4}^*$ and let

$$\underline{\mathbf{v}}_3 = \underline{\mathbf{v}}_3^* = \begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \underline{\mathbf{v}}_4 = \underline{\mathbf{v}}_4^* - \left(\frac{\underline{\mathbf{v}}_4^* \cdot \underline{\mathbf{v}}_3}{\underline{\mathbf{v}}_3 \cdot \underline{\mathbf{v}}_3}\right)\underline{\mathbf{v}}_3 = \begin{pmatrix} -2 \\ -3 \\ 1 \\ 0 \end{pmatrix} - \left(\frac{\begin{pmatrix} -2 \\ -3 \\ 1 \\ 0 \end{pmatrix} . \begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix}}{\begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix} . \begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix}}\right)\begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix} = -\frac{1}{3}\begin{pmatrix} 2 \\ 1 \\ -3 \\ 4 \end{pmatrix},$$

which we can confirm is orthogonal to $\underline{\mathbf{v}}_3$. Finally, with normalization we're done, so

$$\underline{\mathbf{v}}_3 = \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \underline{\mathbf{v}}_4 = \frac{1}{\sqrt{30}}\begin{pmatrix} 2 \\ 1 \\ -3 \\ 4 \end{pmatrix}$$

and so finally, we use $\underline{\mathbf{v}}_i$ as the columns to form the second orthogonal matrix

$$V = \begin{pmatrix} \frac{3}{2\sqrt{5}} & -\frac{1}{2} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{30}} \\ -\frac{1}{2\sqrt{5}} & \frac{1}{2} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{30}} \\ \frac{3}{2\sqrt{5}} & \frac{1}{2} & 0 & -\frac{3}{\sqrt{30}} \\ \frac{1}{2\sqrt{5}} & \frac{1}{2} & \frac{1}{\sqrt{6}} & \frac{4}{\sqrt{30}} \end{pmatrix}$$

We now have the SVD:

$$A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ -2 & 1 & -1 & 0 \\ -1 & 1 & 1 & 1 \end{pmatrix} = \frac{1}{\sqrt{6}}\begin{pmatrix} \sqrt{3} & 1 & \sqrt{2} \\ -\sqrt{3} & 1 & \sqrt{2} \\ 0 & 2 & -\sqrt{2} \end{pmatrix}\begin{pmatrix} \sqrt{10} & 0 & 0 & 0 \\ 0 & \sqrt{6} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} \frac{3}{2\sqrt{5}} & -\frac{1}{2} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{30}} \\ -\frac{1}{2\sqrt{5}} & \frac{1}{2} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{30}} \\ \frac{3}{2\sqrt{5}} & \frac{1}{2} & 0 & -\frac{3}{\sqrt{30}} \\ \frac{1}{2\sqrt{5}} & \frac{1}{2} & \frac{1}{\sqrt{6}} & \frac{4}{\sqrt{30}} \end{pmatrix}^T = U\Sigma V^T.$$