Final Review 2

Problem 1: Consider the matrix:

$$A = \begin{bmatrix} -2 & 3 & 3 \\ 2 & -1 & -1 \\ -7 & 7 & 6 \end{bmatrix}$$

(a) Compute the characteristic polynomial of A.

(b) Compute eigenvalues and eigenvectors of A. What are the algebraic/geometric multiplicities?

(c) Diagonalize the matrix A.

(d) Describe the behavior of the solution to the system $\dot{\boldsymbol{v}}(t) = A\boldsymbol{v}(t)$ as $t \to \infty$, where $\boldsymbol{v}$ is a vector consisting of three functions of t.

(e) Use Cramer's rule to find a solution to the equation $A\boldsymbol{w} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

Solution: (a) The characteristic polynomial is given by:

$$p(\lambda) = \det \begin{bmatrix} -2-\lambda & 3 & 3 \\ 2 & -1-\lambda & -1 \\ -7 & 7 & 6-\lambda \end{bmatrix}$$

By cofactor expansion along the first column, this is given by:

$$p(\lambda) = (-2-\lambda) \cdot \det \begin{bmatrix} -1-\lambda & -1 \\ 7 & 6-\lambda \end{bmatrix} - 2 \cdot \det \begin{bmatrix} 3 & 3 \\ 7 & 6-\lambda \end{bmatrix} + (-7) \cdot \det \begin{bmatrix} 3 & 3 \\ -1-\lambda & -1 \end{bmatrix} =$$

$$= (-2-\lambda)(\lambda^2 - 5\lambda + 1) - 2(-3\lambda - 3) + (-7)(3\lambda) = -\lambda^3 + 3\lambda^2 - 6\lambda + 4$$

(b) In general, there's no formula for the solutions to a cubic polynomial. But in practical cases, you can try and guess a solution (just plug in some easy numbers) which would allow you to chip off a linear factor from the polynomial. In the case at hand, it's easy to see that $\lambda = 1$ is a solution, so the characteristic polynomial factors as:

$$p(\lambda) = (1-\lambda)(\lambda^2 - 2\lambda + 4)$$

Now the second polynomial is quadratic, and its solutions are:

$$\lambda = \frac{2 \pm \sqrt{4-16}}{2} = 1 \pm i\sqrt{3}$$

So the roots of p, i.e. the eigenvalues, are:

$$\lambda_1 = 1, \quad \lambda_2 = 1 + i\sqrt{3}, \quad \lambda_3 = 1 - i\sqrt{3} \tag{1}$$

Remark. *Let's focus on the latter two eigenvalues, since they are conjugate complex numbers. You may want to convert them from Cartesian to polar coordinates, and indeed, then way to do so is to compute:*

$$r = |1 + i\sqrt{3}| = \sqrt{1^2 + \sqrt{3}^2} = \sqrt{4} = 2$$
$$\theta = arccos\left(\frac{1}{2}\right) = \frac{\pi}{3}$$

Therefore, in polar coordinates, we have:

$$1 + i\sqrt{3} = 2e^{\frac{i\pi}{3}} \qquad\qquad 1 - i\sqrt{3} = 2e^{-\frac{i\pi}{3}}$$

Since all the eigenvalues are distinct, they all have algebraic and geometric multiplicity 1. Let's now compute eigenvectors corresponding to the eigenvalues (1):

$$\boldsymbol{v}_1 \in \mathrm{NS}(A - \lambda_1 \cdot I) = \mathrm{NS}\begin{bmatrix} -3 & 3 & 3 \\ 2 & -2 & -1 \\ -7 & 7 & 5 \end{bmatrix}$$

To compute the nullspace to the right, let's apply Gauss-Jordan decomposition:

$$\begin{bmatrix} -3 & 3 & 3 \\ 2 & -2 & -1 \\ -7 & 7 & 5 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

and therefore:

$$\boldsymbol{v}_1 = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{where} \quad \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0, \quad \text{i.e.} \quad \begin{cases} x - y = 0 \\ z = 0 \end{cases}$$

The pivot variables are x and z, and the free variables are y. So a choice of eigenvector is obtained by setting the free variable equal to 1:

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

As for the second eigenvalue, we have:

$$\boldsymbol{v}_2 \in \mathrm{NS}(A - \lambda_2 \cdot I) = \mathrm{NS}\begin{bmatrix} -3 - i\sqrt{3} & 3 & 3 \\ 2 & -2 - i\sqrt{3} & -1 \\ -7 & 7 & 56 - i\sqrt{3} \end{bmatrix}$$

Since it involves algebra with complex numbers, let's do every step of Gauss-Jordan decomposition. First, we must subtract:

$$\frac{-7}{-3 - i\sqrt{3}} = \frac{-7(-3 + i\sqrt{3})}{3^2 + \sqrt{3}^2} = \frac{21 - i7\sqrt{3}}{12}$$

times the first row from the third row:

$$\begin{bmatrix} -3 - i\sqrt{3} & 3 & 3 \\ 2 & -2 - i\sqrt{3} & -1 \\ -7 & 7 & 5 - i\sqrt{3} \end{bmatrix} \rightsquigarrow \begin{bmatrix} -3 - i\sqrt{3} & 3 & 3 \\ 2 & -2 - i\sqrt{3} & -1 \\ 0 & 7 - \frac{21 - i7\sqrt{3}}{4} & 5 - i\sqrt{3} - \frac{21 - i7\sqrt{3}}{4} \end{bmatrix} =$$

$$= \begin{bmatrix} -3-i\sqrt{3} & 3 & 3 \\ 2 & -2-i\sqrt{3} & -1 \\ 0 & \frac{7}{4}(1+i\sqrt{3}) & \frac{-1+i3\sqrt{3}}{4} \end{bmatrix}$$

Then we must subtract:

$$\frac{2}{-3-i\sqrt{3}} = \frac{2(-3+i\sqrt{3})}{3^2+\sqrt{3}^2} = \frac{-6+i2\sqrt{3}}{12}$$

times the first row from the second row:

$$\begin{bmatrix} -3-i\sqrt{3} & 3 & 3 \\ 2 & -2-i\sqrt{3} & -1 \\ 0 & \frac{7}{4}(1+i\sqrt{3}) & \frac{-1+i3\sqrt{3}}{4} \end{bmatrix} \rightsquigarrow \begin{bmatrix} -3-i\sqrt{3} & 3 & 3 \\ 0 & -2-i\sqrt{3}-\frac{-6+i2\sqrt{3}}{4} & -1-\frac{-6+i2\sqrt{3}}{4} \\ 0 & \frac{7}{4}(1+i\sqrt{3}) & \frac{-1+i3\sqrt{3}}{4} \end{bmatrix} =$$

$$= \begin{bmatrix} -3-i\sqrt{3} & 3 & 3 \\ 0 & \frac{-2-i6\sqrt{3}}{4} & \frac{2-i2\sqrt{3}}{4} \\ 0 & \frac{7}{4}(1+i\sqrt{3}) & \frac{-1+i3\sqrt{3}}{4} \end{bmatrix}$$

Finally, we must subtract:

$$\frac{7(1+i\sqrt{3})}{-2-i6\sqrt{3}} = \frac{7(1+i\sqrt{3})(-2+i6\sqrt{3})}{(-2)^2+6^2\sqrt{3}^2} = \frac{7(-2-i2\sqrt{3}+i6\sqrt{3}-6\cdot 3)}{112} = \frac{-5+i\sqrt{3}}{4}$$

times the second row from the third row. Since $\frac{2-i2\sqrt{3}}{4} \cdot \frac{-5+i\sqrt{3}}{4} = \frac{-1+i3\sqrt{3}}{4}$, we obtain:

$$\begin{bmatrix} -3-i\sqrt{3} & 3 & 3 \\ 0 & \frac{-2-i6\sqrt{3}}{4} & \frac{2-i2\sqrt{3}}{4} \\ 0 & \frac{3+i7\sqrt{3}}{4} & \frac{3+i3\sqrt{3}}{4} \end{bmatrix} \rightsquigarrow \begin{bmatrix} -3-i\sqrt{3} & 3 & 3 \\ 0 & \frac{-1-i3\sqrt{3}}{2} & \frac{1-i\sqrt{3}}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

To make all the pivots 1, we need to multiply the first and second row, respectively, by:

$$\frac{1}{-3-i\sqrt{3}} = \frac{-3+i\sqrt{3}}{12} \qquad \text{and} \qquad \frac{2}{-1-i3\sqrt{3}} = \frac{-1+i3\sqrt{3}}{14}$$

and we get:

$$\begin{bmatrix} -3-i\sqrt{3} & 3 & 3 \\ 0 & \frac{-1-i3\sqrt{3}}{2} & \frac{1-i\sqrt{3}}{2} \\ 0 & 0 & 0 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & \frac{-3+i\sqrt{3}}{4} & \frac{-3+i\sqrt{3}}{4} \\ 0 & 1 & \frac{1-i\sqrt{3}}{2}\cdot\frac{-1+i3\sqrt{3}}{14} \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & \frac{-3+i\sqrt{3}}{4} & \frac{-3+i\sqrt{3}}{4} \\ 0 & 1 & \frac{2+i\sqrt{3}}{7} \\ 0 & 0 & 0 \end{bmatrix}$$

Finally, we need to subtract $\frac{-3+i\sqrt{3}}{4}$ times the second row from the first row:

$$\begin{bmatrix} 1 & \frac{-3+i\sqrt{3}}{4} & \frac{-3+i\sqrt{3}}{4} \\ 0 & 1 & \frac{2+i\sqrt{3}}{7} \\ 0 & 0 & 0 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & \frac{-3+i\sqrt{3}}{4} - \frac{2+i\sqrt{3}}{7}\frac{-3+i\sqrt{3}}{4} \\ 0 & 1 & \frac{2+i\sqrt{3}}{7} \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \frac{-3+i2\sqrt{3}}{7} \\ 0 & 1 & \frac{2+i\sqrt{3}}{7} \\ 0 & 0 & 0 \end{bmatrix}$$

Therefore:

$$\boldsymbol{v}_2 = \begin{bmatrix} 1 & 0 & \frac{-3+i2\sqrt{3}}{7} \\ 0 & 1 & \frac{2+i\sqrt{3}}{7} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0, \quad \text{i.e.} \quad \begin{cases} x + z\frac{-3+i2\sqrt{3}}{7} = 0 \\ y + z\frac{2+i\sqrt{3}}{7} = 0 \end{cases}$$

The pivot variables are x and y, and the free variables are z. So a choice of eigenvector is obtained by setting the free variable equal to 1:

$$\boldsymbol{v}_2 = \begin{bmatrix} \frac{3-i2\sqrt{3}}{7} \\ \frac{-2-i\sqrt{3}}{7} \\ 1 \end{bmatrix}$$

Since the eigenvalue λ_3 is the conjugate of the eigenvalue λ_2, the corresponding eigenvector $\boldsymbol{v}_3$ is the conjugate of the eigenvector $\boldsymbol{v}_2$ found above (this is a general feature which only requires A to have real entries):

$$\boldsymbol{v}_3 = \begin{bmatrix} \frac{3+i2\sqrt{3}}{7} \\ \frac{-2+i\sqrt{3}}{7} \\ 1 \end{bmatrix}$$

(c) The diagonalization of A is $A = VDV^{-1}$, where:

$$V = \left[\ \boldsymbol{v}_1 \mid \boldsymbol{v}_2 \mid \boldsymbol{v}_3\ \right] = \begin{bmatrix} 1 & \frac{3-i2\sqrt{3}}{7} & \frac{3+i2\sqrt{3}}{7} \\ 1 & \frac{-2-i\sqrt{3}}{7} & \frac{-2+i\sqrt{3}}{7} \\ 0 & 1 & 1 \end{bmatrix}$$

and:

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1+i\sqrt{3} & 0 \\ 0 & 0 & 1-i\sqrt{3} \end{bmatrix}$$

(d) The solutions to the system in question are of the form:

$$\boldsymbol{v}(t) = e^{At}\boldsymbol{c} = Ve^{Dt}V^{-1}\boldsymbol{c} = V \begin{bmatrix} e^{\lambda_1 t} & 0 & 0 \\ 0 & e^{\lambda_2 t} & 0 \\ 0 & 0 & e^{\lambda_3 t} \end{bmatrix} V^{-1}\boldsymbol{c} = V \begin{bmatrix} e^{t} & 0 & 0 \\ 0 & e^{t+it\sqrt{3}} & 0 \\ 0 & 0 & e^{t-it\sqrt{3}} \end{bmatrix} V^{-1}\boldsymbol{c}$$

where $\boldsymbol{c}$ is an arbitrary vector with constant entries. Leaving aside the particular value of these constants, the behavior of $\boldsymbol{v}(t)$ as $t \to \infty$ is determined by the rate of growth of the functions:

$$e^t, e^{t+it\sqrt{3}}, e^{t-it\sqrt{3}}$$

The former of these is well-known to you: it is an exponential. The second and third are equal to:

$$e^t \cdot e^{\pm it\sqrt{3}} = e^t\Big(\cos(t\sqrt{3}) \pm i \cdot \sin(t\sqrt{3})\Big)$$

As $t \to \infty$, the function above tends to ∞ but oscillates between $+$ and $-$ infinity in accordance with the way sine and cosine oscillate between positive and negative values.

(e) The determinant of A is 4, since it matches the value of the characteristic polynomial at $\lambda = 0$. Cramer's rule states that the solution is given by:

$$\boldsymbol{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

where:

$$w_1 = \frac{1}{4}\det\begin{bmatrix}1 & 3 & 3\\ 0 & -1 & -1\\ 0 & 7 & 6\end{bmatrix} = \frac{1}{4}\det\begin{bmatrix}-1 & -1\\ 7 & 6\end{bmatrix} = \frac{1}{4}$$

$$w_2 = \frac{1}{4}\det\begin{bmatrix}-2 & 1 & 3\\ 2 & 0 & -1\\ -7 & 0 & 6\end{bmatrix} = -\frac{1}{4}\det\begin{bmatrix}2 & -1\\ -7 & 6\end{bmatrix} = -\frac{5}{4}$$

$$w_3 = \frac{1}{4}\det\begin{bmatrix}-2 & 3 & 1\\ 2 & -1 & 0\\ -7 & 7 & 0\end{bmatrix} = \frac{1}{4}\det\begin{bmatrix}2 & -1\\ -7 & 7\end{bmatrix} = \frac{7}{4}$$

(the determinants above were computed by cofactor expansion along the column with many zeroes).

Problem 2: Consider the matrix:

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

(a) Compute the SVD of A.

(b) Consider the symmetric matrices A^TA and AA^T. Are they positive definite, positive semi-definite, or neither? What are the corresponding energy functions of these symmetric matrices?

(c) Find numbers a and b for which the quantity:

$$(a\sqrt{2}-1)^2 + (a+b-1)^2 + (-b\sqrt{2}-1)^2$$

is minimal.

Solution: (a) We need to compute the singular vectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ (which form an orthonormal basis of $\mathbb{R}^2$), the singular values σ_1, σ_2 and the singular vectors $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{u}_3$ (which form an orthonormal basis of $\mathbb{R}^3$). To compute the $\boldsymbol{v}$'s and the σ's, we need to diagonalize the symmetric matrix:

$$A^TA = \begin{bmatrix}3 & 1\\ 1 & 3\end{bmatrix}$$

Its characteristic polynomial is:

$$p(\lambda) = (3-\lambda)^2 - 1 = \lambda^2 - 6\lambda - 8 = (\lambda-4)(\lambda-2)$$

hence its eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = 4$. The singular values are the square roots of these, so we have $\sigma_1 = \sqrt{2}$ and $\sigma_2 = 2$. The vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are eigenvectors of A^TA, and the usual procedure (see the previous problem for the calculation of eigenvectors) yields:

$$\boldsymbol{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\ -1\end{bmatrix} \qquad \text{and} \qquad \boldsymbol{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\ 1\end{bmatrix}$$

(the reason for the scalar factor $\frac{1}{\sqrt{2}}$ in front is that you want $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ to have length 1, in order to give an orthonormal basis). With the $\boldsymbol{v}$'s and the σ's in hand, we can define $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ by the equations:

$$A\boldsymbol{v}_1 = \sigma_1 \boldsymbol{u}_1 \quad \Rightarrow \quad \boldsymbol{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{2} & 0 \\ 1 & 1 \\ 0 & -\sqrt{2} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$A\boldsymbol{v}_2 = \sigma_2 \boldsymbol{u}_2 \quad \Rightarrow \quad \boldsymbol{u}_2 = \frac{1}{2} \begin{bmatrix} \sqrt{2} & 0 \\ 1 & 1 \\ 0 & -\sqrt{2} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ \sqrt{2} \\ -1 \end{bmatrix}$$

Now we just need to compute $\boldsymbol{u}_3$, and it's given by the property that it forms as orthonormal basis together with $\boldsymbol{u}_1$, $\boldsymbol{u}_2$. The standard way to do this is Gram-Schmidt. Consider any basis:

$$\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{a}$$

where $\boldsymbol{a}$ is some arbitrary simple vector that is independent from $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, say:

$$\boldsymbol{a} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

To modify the basis $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{a}$ in such a way as to make it orthonormal, we need to replace:

$$\boldsymbol{a} \rightsquigarrow \boldsymbol{b} = \boldsymbol{a} - \mathrm{proj}_{\boldsymbol{u}_1} \boldsymbol{a} - \mathrm{proj}_{\boldsymbol{u}_2} \boldsymbol{a} = \boldsymbol{a} - \boldsymbol{u}_1(\boldsymbol{u}_1^T \boldsymbol{a}) - \boldsymbol{u}_2(\boldsymbol{u}_2^T \boldsymbol{a}) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{4} \begin{bmatrix} 1 \\ \sqrt{2} \\ -1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 \\ -\sqrt{2} \\ -1 \end{bmatrix}$$

The vector $\boldsymbol{b}$ thus defined is orthogonal to $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ by construction, and its length is $\frac{1}{2}$. Therefore, we must renormalize it to have length 1, and the correct choice for $\boldsymbol{u}_3$ is:

$$\boldsymbol{u}_3 = 2\boldsymbol{b} = \frac{1}{2} \begin{bmatrix} 1 \\ -\sqrt{2} \\ -1 \end{bmatrix}$$

We conclude that the SVD of A is:

$$A = U \Sigma V^T$$

where:

$$U = \left[\ \boldsymbol{u}_1 \ \middle| \ \boldsymbol{u}_2 \ \middle| \ \boldsymbol{u}_3 \ \right], \qquad \Sigma = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}, \qquad V = \left[\ \boldsymbol{v}_1 \ \middle| \ \boldsymbol{v}_2 \ \right]$$

with $\boldsymbol{v}_1, \boldsymbol{v}_2$, $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ computed as above.

As we have seen, $A^T A$ has eigenvalues 2 and 4. Since they are both positive, we conclude that $A^T A$ is positive definite. Meanwhile, we have:

$$AA^T = \begin{bmatrix} 2 & \sqrt{2} & 0 \\ \sqrt{2} & 2 & -\sqrt{2} \\ 0 & -\sqrt{2} & 2 \end{bmatrix}$$

The positive eigenvalues of AA^T are the same as those of A^TA, namely 2 and 4. However, the 3×3 matrix AA^T has another eigenvalues, which must be a 0. Therefore, the matrix AA^T is only positive semidefinite, because its eigenvalues are non-negative (but it cannot be a positive definite matrix because it has a 0 eigenvalue).

The energy functions are:

$$A^TA \quad \rightsquigarrow \quad \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 3x^2 + 2xy + 3y^2$$

$$AA^T \quad \rightsquigarrow \quad \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 2 & \sqrt{2} & 0 \\ \sqrt{2} & 2 & -\sqrt{2} \\ 0 & -\sqrt{2} & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2x^2 + 2y^2 + 2z^2 + 2\sqrt{2}xy - 2\sqrt{2}yz$$

Note that the quantity $f(x, y) = 3x^2 + 2xy + 3y^2$ is always positive (unless $x = y = 0$) because the matrix A^TA is positive definite. However, the quantity $g(x, y, z) = 2x^2+2y^2+2z^2+2\sqrt{2}xy-2\sqrt{2}yz$ is only non-negative, and indeed it does have zero values. One situation when this happens is when:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \boldsymbol{u}_3 = \frac{1}{2} \begin{bmatrix} 1 \\ -\sqrt{2} \\ -1 \end{bmatrix}$$

For the values of x, y, z in the equation above, you can see that $g(x, y, z)$ by direct calculation. But probably a more conceptual way to see this is right from the SVD, which entails the fact that:

$$A^T\boldsymbol{u}_3 = 0 \quad \Rightarrow \quad AA^T\boldsymbol{u}_3 = 0 \quad \Rightarrow \quad \boldsymbol{u}_3^T(AA^T)\boldsymbol{u}_3 = 0$$

(c) The problem can be phrased as a least squares approximation problem. Specifically, the sum of squares should always have you think about the length squared of a vector. A little thought shows that the way to do this is:

$$(a\sqrt{2} - 1)^2 + (a + b - 1)^2 + (-b\sqrt{2} - 1)^2 = \left\| A \begin{bmatrix} a \\ b \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\|^2$$

Therefore, we need to find that vector $\boldsymbol{v} = \begin{bmatrix} a \\ b \end{bmatrix}$ such that $A\boldsymbol{v}$ is as close as possible to $\boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

The principle of least squares tells us we must have:

$$A\boldsymbol{v} = \text{proj}_{C(A)}\boldsymbol{b} = A(A^TA)^{-1}A^T\boldsymbol{b}$$

so we have to choose:

$$\boldsymbol{v} = (A^TA)^{-1}A^T\boldsymbol{b} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 1 & 0 \\ 0 & 1 & -\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 + 2\sqrt{2} \\ 1 - 2\sqrt{2} \end{bmatrix}$$

Note that we could have also written $v = A^+\boldsymbol{b}$, where A^+ is the pseudo-inverse of A. This is because if we plug the formula for the SVD of A, we have:

$$(A^TA)^{-1}A^T = (V\Sigma^T\Sigma V^{-1})^{-1}V\Sigma^TU^T = V(\Sigma^T\Sigma)^{-1}\Sigma^TU^T = V\Sigma^+U^T = A^+$$

The next-to-last identity is simply direct computation:

$$(\Sigma^T \Sigma)^{-1}\Sigma^T = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}^{-1} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & 0 \\ 0 & \frac{1}{\sigma_2} & 0 \end{bmatrix}$$

Problem 3: Consider two random variables X and Y, which take values:

$$\{x = 1 \text{ and } y = 1\} \quad \text{with probability } \frac{2}{3}$$
$$\{x = 0 \text{ and } y = 3\} \quad \text{with probability } \frac{1}{3}$$

(a) Find linear combinations of x and y which are uncorrelated (i.e. have covariance 0).

(b) Find a general formula, in terms of matrices and vectors, for the covariance of any two linear combinations of these random variables: $ax + by$ and $a'x + b'y$ (where a, b, a', b' are numbers).

Solution: (a) We start by converting the problem into linear algebra. The first step is to put the two random variables into a random vector:

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \end{bmatrix}$$

Then the given input tells us that this random vector takes the value:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{with probability } \frac{2}{3}$$
$$\begin{bmatrix} 0 \\ 3 \end{bmatrix} \quad \text{with probability } \frac{1}{3}$$

Therefore, the mean (or expected value) of the random vector is:

$$\boldsymbol{\mu} = \frac{2}{3}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{5}{3} \end{bmatrix}$$

By definition, the covariance matrix is the following expectation value:

$$K = E[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})^T] = \frac{2}{3}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{2}{3} \\ \frac{5}{3} \end{bmatrix}\right)\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{2}{3} \\ \frac{5}{3} \end{bmatrix}\right)^T + \frac{1}{3}\left(\begin{bmatrix} 0 \\ 3 \end{bmatrix} - \begin{bmatrix} \frac{2}{3} \\ \frac{5}{3} \end{bmatrix}\right)\left(\begin{bmatrix} 0 \\ 3 \end{bmatrix} - \begin{bmatrix} \frac{2}{3} \\ \frac{5}{3} \end{bmatrix}\right)^T =$$

$$= \frac{2}{3}\begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix}\begin{bmatrix} \frac{1}{3} & -\frac{2}{3} \end{bmatrix} + \frac{1}{3}\begin{bmatrix} -\frac{2}{3} \\ \frac{4}{3} \end{bmatrix}\begin{bmatrix} -\frac{2}{3} & \frac{4}{3} \end{bmatrix} = \frac{2}{9}\begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} + \frac{1}{9}\begin{bmatrix} 4 & -8 \\ -8 & 16 \end{bmatrix} = \frac{2}{3}\begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} \qquad (2)$$

The key step now is to diagonalize the covariance matrix. I won't bore you with the details, because it's done just like in problem 1, but the answer is:

$$K = Q\begin{bmatrix} 0 & 0 \\ 0 & \frac{10}{3} \end{bmatrix}Q^T \quad \text{where} \quad Q = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}$$

So if we consider the random vector $\boldsymbol{Y} = Q^T \boldsymbol{X}$, then the covariance matrix of $\boldsymbol{Y}$ is:

$$E[(\boldsymbol{Y} - E[\boldsymbol{Y}])(\boldsymbol{Y} - E[\boldsymbol{Y}])^T] = Q^T K Q = \begin{bmatrix} 0 & 0 \\ 0 & \frac{10}{3} \end{bmatrix} \tag{3}$$

Therefore, the entries of $\boldsymbol{Y}$ are uncorrelated, and they are the linear combinations we are looking for. Explicitly:

$$\boldsymbol{Y} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{2x+y}{\sqrt{5}} \\ \frac{-x+2y}{\sqrt{5}} \end{bmatrix}$$

so our analysis shows that the linear combinations $\frac{2x+y}{\sqrt{5}}$ and $\frac{-x+2y}{\sqrt{5}}$ are indeed uncorrelated. Moreover, the variances of these linear combinations can be read off from the diagonal entries of (3):

$$\frac{2x+y}{\sqrt{5}} \text{ has covariance } 0$$
$$\frac{-x+2y}{\sqrt{5}} \text{ has covariance } \frac{10}{3}$$

The first of these statements says that the random variable $2x + y$ is a constant. And indeed, this is the case because $2x + y$ is equal to 3 in either of the two cases prescribed by the probability distribution in the problem.

(b) The covariance of any two random variables is linear, so:

$$K_{ax+by,a'x+b'y} = aa'K_{xx} + ab'K_{xy} + ba'K_{yx} + bb'K_{yy}$$

If we recall that the covariance matrix is:

$$K = \begin{bmatrix} K_{xx} & K_{xy} \\ K_{yx} & K_{yy} \end{bmatrix}$$

then we conclude the formula:

$$K_{ax+by,a'x+b'y} = \begin{bmatrix} a & b \end{bmatrix} K \begin{bmatrix} a' \\ b' \end{bmatrix}$$

You can get a number out of this by plugging any numbers you want for a, b, a', b', and the matrix K from (2).