The Austalian National University
School of Computing
Liang Zhang

Semester 2, 2021
Tutorial 9

COMP3670/6670: Introduction to Machine Learning

Question 1 **Laplace Expansion**

Given the following matrix $\mathbf{A}$,

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 0 & 3 \\ 2 & -1 & 5 & -3 \\ 0 & 4 & 0 & 10 \\ 1 & 3 & 1 & 4 \end{bmatrix}$$

Verify that $\det \mathbf{A} = -16$ by using Laplace expansion.

Solution. We choose to expand along the third column, as it contains two zeros.

$$\begin{aligned}
&\det \mathbf{A} \\
&= 0 \cdot \det \begin{bmatrix} 2 & -1 & -3 \\ 0 & 4 & 10 \\ 1 & 3 & 4 \end{bmatrix} - 5 \cdot \det \begin{bmatrix} 1 & 4 & 3 \\ 0 & 4 & 10 \\ 1 & 3 & 4 \end{bmatrix} + 0 \cdot \det \begin{bmatrix} 1 & 4 & 3 \\ 2 & -1 & -3 \\ 1 & 3 & 4 \end{bmatrix} - 1 \cdot \det \begin{bmatrix} 1 & 4 & 3 \\ 2 & -1 & -3 \\ 0 & 4 & 10 \end{bmatrix} \\
&= -5 \cdot \det \begin{bmatrix} 1 & 4 & 3 \\ 0 & 4 & 10 \\ 1 & 3 & 4 \end{bmatrix} - 1 \cdot \det \begin{bmatrix} 1 & 4 & 3 \\ 2 & -1 & -3 \\ 0 & 4 & 10 \end{bmatrix}
\end{aligned}$$

For the first matrix, we apply Laplace expansion again along the first column, as it contains a zero, and the other coefficients are small.

$$\begin{aligned}
&= \det \begin{bmatrix} 1 & 4 & 3 \\ 0 & 4 & 10 \\ 1 & 3 & 4 \end{bmatrix} \\
&= 1 \cdot \det \begin{bmatrix} 4 & 10 \\ 3 & 4 \end{bmatrix} - 0 \cdot \det \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix} + 1 \cdot \det \begin{bmatrix} 4 & 3 \\ 4 & 10 \end{bmatrix} \\
&= \det \begin{bmatrix} 4 & 10 \\ 3 & 4 \end{bmatrix} + \det \begin{bmatrix} 4 & 3 \\ 4 & 10 \end{bmatrix} \\
&= (4 \cdot 4 - 3 \cdot 10) + (4 \cdot 10 - 3 \cdot 4) = (16 - 30) + (40 - 12) = -14 + 28 = 14
\end{aligned}$$

For the second matrix, we apply Laplace expansion along the first column, as it contains a zero, and the coefficients are small.

$$\begin{aligned}
&\det \begin{bmatrix} 1 & 4 & 3 \\ 2 & -1 & -3 \\ 0 & 4 & 10 \end{bmatrix} \\
&= 1 \cdot \begin{bmatrix} -1 & -3 \\ 4 & 10 \end{bmatrix} - 2 \cdot \begin{bmatrix} 4 & 3 \\ 4 & 10 \end{bmatrix} + 0 \cdot \begin{bmatrix} 4 & 3 \\ -1 & -3 \end{bmatrix} \\
&= 1 \cdot \begin{bmatrix} -1 & -3 \\ 4 & 10 \end{bmatrix} - 2 \cdot \begin{bmatrix} 4 & 3 \\ 4 & 10 \end{bmatrix} \\
&= (-1 \cdot 10 - (-3) \cdot 4) - 2(4 \cdot 10 - 3 \cdot 4) \\
&= (-10 + 12) - 2(40 - 12) \\
&= 2 - 56 = -54
\end{aligned}$$

We can now combine the results together.

$$\det \mathbf{A} = -5 \cdot 14 - 1 \cdot (-54) = -70 + 54 = -16$$

Question 2 **Upper Triangular Matrix**

An upper triangular matrix is any square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ where all the elements below the main diagonal are zero (that is, for all $i > j$, $A_{ij} = 0$.)

1. Prove that the set of eigenvalues of any upper triangular matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is the same as the set of diagonal elements $\{A_{11}, A_{22}, \ldots, A_{nn}\}$.

   (Hint: Use the fact that the determinant of an upper triangular matrix is the product of it's diagonal elements, which you will prove in the assignment.)

   Solution. The set of all eigenvalues is given by the solutions of the characteristic equation $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$. Note that if $\mathbf{A}$ is upper triangular, then so is $\mathbf{A} - \lambda \mathbf{I}$, as we are just subtracting λ from each diagonal element. Applying the property from the assignment, we have

   $$\det(\mathbf{A} - \lambda \mathbf{I}) = \prod_{i=1}^{n} (\mathbf{A} - \lambda \mathbf{I})_{ii} = \prod_{i=1}^{n} (A_{ii} - \lambda)$$

   Hence, we have

   $$(A_{11} - \lambda)(A_{22} - \lambda) \ldots (A_{nn} - \lambda) = 0$$

   which has solutions

   $$\lambda = A_{11} \text{ or } \lambda = A_{22} \text{ or } \ldots \text{ or } \lambda = A_{nn}$$

   which are just the diagonal elements of $\mathbf{A}$, as required.

2. Prove that the set of all $n \times n$ upper triangular matrices are closed under matrix multiplication.

   Solution. Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ be upper triangular matricies. We want to prove that $\mathbf{AB}$ is also upper triangular, that is, for all $i > j, (\mathbf{AB})_{ij} = 0$. Let i, j be indexes in the range 1 to n with the constraint that $i > j$.

   $$(\mathbf{AB})_{ij}$$
   $$= \sum_{m=1}^{n} \mathbf{A}_{im} \mathbf{B}_{mj}$$

   Now, if $i > m$, then $\mathbf{A}_{im} = 0$, so we can ignore these terms.

   $$= \sum_{m=1, i \leq m}^{n} \mathbf{A}_{im} \mathbf{B}_{mj}$$

   If $m > j$, then $\mathbf{B}_{mj} = 0$, so we can also ignore these terms.

   $$= \sum_{m=1, i \leq m, m \leq j}^{n} \mathbf{A}_{im} \mathbf{B}_{mj} = \sum_{m=1, i \leq m \leq j}^{n} \mathbf{A}_{im} \mathbf{B}_{mj}$$

   But now we are summing over all values of m such that $i \leq m \leq j$, but since by construction $i > j$, there cannot exist any such m. Hence, we are summing over no values, and the sum is zero.

   Hence, for all $i > j, (\mathbf{AB})_{ij} = 0$, as required.

Question 3 **Fast Matrix Exponentiation**

1. Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be a diagonal matrix. Prove that for any integer $k \geq 1$ that

$$(\mathbf{A}^k)_{ij} = (\mathbf{A}_{ij})^k$$

Solution. We proceed via induction. Trivial for $k = 1$. Assume true for some k. Now,

$$\begin{aligned}
&\mathbf{A}_{ij}^{k+1} \\
&= (\mathbf{A}^k\mathbf{A})_{ij} \\
&= \sum_{m=1}^{k} \mathbf{A}_{im}^k \mathbf{A}_{mj} \\
&= \sum_{m=1}^{k} (\mathbf{A}_{im})^k \mathbf{A}_{mj}
\end{aligned}$$

Since $\mathbf{A}$ is a diagonal matrix, $\mathbf{A}_{im}$ is zero whenever $i \neq m$, so we can ignore those terms.

$$= (\mathbf{A}_{ii})^k \mathbf{A}_{ij}$$

Now, if $i \neq j$, then $\mathbf{A}_{ij} = 0$, and so $(\mathbf{A}_{ii})^k \mathbf{A}_{ij} = (\mathbf{A}_{ij})^{k+1}$ and we are done. So, assume that $i = j$.

$$= (\mathbf{A}_{ii})^k \mathbf{A}_{ii} = (\mathbf{A}_{ii})^{k+1}$$

as required.

2. Let $\mathbf{B} \in \mathbb{R}^{n \times n}$ be a diagonalizable matrix, with an invertible matrix $\mathbf{Q} \in \mathbb{R}^{n \times n}$ and a diagonal matrix $\mathbf{E} \in \mathbb{R}^{n \times n}$ satisfying $\mathbf{B} = \mathbf{Q}\mathbf{E}\mathbf{Q}^{-1}$. Prove that for any integer $k \geq 1$, that

$$\mathbf{B}^k = \mathbf{Q}\mathbf{E}^k\mathbf{Q}^{-1}$$

(Hint: Use induction on k. See appendix if you are unfamiliar with an induction proof.)

Why does this allow $\mathbf{B}^k$ to be computed quickly?

Solution. Proof by induction on k. Trivial for $k = 1$, as it is the same as the given property. Assume true for some k. For $k + 1$, we have

$$\begin{aligned}
\mathbf{B}^{k+1} &= \mathbf{B}\mathbf{B}^k \\
&= (\mathbf{Q}\mathbf{E}\mathbf{Q}^{-1})(\mathbf{Q}\mathbf{E}^k\mathbf{Q}^{-1}) \\
&= \mathbf{Q}\mathbf{E}(\mathbf{Q}^{-1}\mathbf{Q})\mathbf{E}^k\mathbf{Q}^{-1} \\
&= \mathbf{Q}\mathbf{E}\mathbf{I}\mathbf{E}^k\mathbf{Q}^{-1} \\
&= \mathbf{Q}\mathbf{E}\mathbf{E}^k\mathbf{Q}^{-1} \\
&= \mathbf{Q}\mathbf{E}^{k+1}\mathbf{Q}^{-1}
\end{aligned}$$

as required.
This allows us to compute $\mathbf{B}^k$ quickly, as $\mathbf{E}^k$ is fast to compute, since $\mathbf{E}$ is a diagonal matrix.

Question 4 **Computing Eigenvalues and Eigenvectors**

Given the following matrix

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

1. Compute the determinant of $\mathbf{A}$.

   Solution. We use Sarrus' rule.

$$\det \mathbf{A} = 2 \cdot 1 \cdot 2 + 0 + 0 - 1 \cdot 1 \cdot 1 - 0 - 0 = 4 - 1 = 3$$

2. What is the characteristic equation of this matrix?

   Solution. We form the matrix $\mathbf{A} - \lambda\mathbf{I}$, and compute it's determinate to obtain the characteristic equation.

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{bmatrix} 2-\lambda & 0 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 0 & 2-\lambda \end{bmatrix}$$

   Applying Sarrus' rule,

$$\begin{aligned} \det(\mathbf{A} - \lambda\mathbf{I}) &= (2-\lambda)(1-\lambda)(2-\lambda) + 0 + 0 - 1 \cdot (1-\lambda) \cdot 1 - 0 - 0 \\ &= (2-\lambda)^2(1-\lambda) - (1-\lambda) \\ &= (1-\lambda)\left((2-\lambda)^2 - 1\right) \end{aligned}$$

   We could have also achieved the same result via Gaussian elimination to a upper triangular matrix, and then applied the previous proven property.

3. Find the eigenvalues, and their algebraic multiplicity.

   Solution. Solving $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, we obtain,

$$1 - \lambda = 0 \text{ or } (2-\lambda)^2 - 1 = 0$$

   In the first case, we have $\lambda = 1$. In the second case, we have $(2-\lambda)^2 = 1$, so $2 - \lambda = \pm 1$. This means that we have $2 - \lambda = 1 \Rightarrow \lambda = 1$, or that $2 - \lambda = -1 \Rightarrow \lambda = 3$.
   Hence, we have that $\lambda = 1$ is an eigenvalue with algebraic multiplicity of 2, and $\lambda = 3$ is an eigenvalue with algebraic multiplicity of 1.

4. For each eigenvalue, compute the corresponding eigenspaces.

   Solution. For $\lambda = 1$ we need to find all $\mathbf{x}$ that solves the equation $\mathbf{A}\mathbf{x} = \mathbf{x}$. We form the homogeneous system $(\mathbf{A} - \mathbf{I})\mathbf{x} = \mathbf{0}$, and solve

$$\left[\begin{array}{ccc|c} 2-1 & 0 & 1 & 0 \\ 1 & 1-1 & 1 & 0 \\ 1 & 0 & 2-1 & 0 \end{array}\right] = \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

   Hence we have that $x_1 = -x_3$ and that both x_2 and x_3 are free variables. The solution set is

$$E_1 = \left\{ \begin{bmatrix} -\beta \\ \alpha \\ \beta \end{bmatrix} : \alpha, \beta \in \mathbb{R} \right\} = \text{span}[\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}]$$

   For $\lambda = 3$ we need to find all $\mathbf{x}$ that solves the equation $\mathbf{A}\mathbf{x} = 3\mathbf{x}$. We form the homogeneous system $(\mathbf{A} - 3\mathbf{I})\mathbf{x} = \mathbf{0}$, and solve

$$\left[\begin{array}{ccc|c} 2-3 & 0 & 1 & 0 \\ 1 & 1-3 & 1 & 0 \\ 1 & 0 & 2-3 & 0 \end{array}\right] = \left[\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 1 & -2 & 1 & 0 \\ 1 & 0 & -1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

   Hence we obtain $x_1 = x_3$ and $x_2 = x_3$, providing the solution set

$$E_3 = \left\{ \begin{bmatrix} \alpha \\ \alpha \\ \alpha \end{bmatrix} : \alpha \in \mathbb{R} \right\} = \text{span}[\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}]$$

5. Show that $\mathbf{A}$ is diagonalizable.

Solution. We need to show that $\mathbf{A}$ is non-defective, that is, that the set of all eigenvectors (the eigenspectrum) spans $\mathbb{R}^3$. We can demonstate that by finding three linearly independent eigenvectors. From Question 3.4 we have a description of the eigenspaces in terms of the spans of basis vectors. By combining the eigenspaces, if the three basis vectors span $\mathbb{R}^3$, then $\mathbf{A}$ is diagonalizable. Using Gaussian elimination, we can show this to be true.

$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

6. Find an invertible matrix $\mathbf{P} \in \mathbb{R}^{n\times n}$ and a diagonal matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$ such that

$$\mathbf{A} = \mathbf{PDP}^{-1}$$

Solution. We can read off the eigenvectors to form the columns of $\mathbf{P}$, and the corresponding eigenvalues (in the same order) to form the diagonal elements of $\mathbf{D}$.

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

$$\mathbf{D} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The inverse of $\mathbf{P}$ can be found via the standard row reduction algorithm.

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} \\ 0 & 0 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} \end{array}\right]$$

Hence,

$$\mathbf{P}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 1 \\ -1 & 2 & -1 \\ -1 & 0 & 1 \end{bmatrix}$$

It can be then verified by standard matrix multiplication that $\mathbf{A} = \mathbf{PDP}^{-1}$.

7. Using Question 3, give a closed form for $\mathbf{A}^n$ for any $n \geq 0$.

Solution. Since $\mathbf{A} = \mathbf{PDP}^{-1}$, we can use Question to compute $\mathbf{A}^n$.

$$\begin{aligned}
\mathbf{A}^n &= \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^n \left(\frac{1}{2}\begin{bmatrix} 1 & 0 & 1 \\ -1 & 2 & -1 \\ -1 & 0 & 1 \end{bmatrix} \right) \\
&= \frac{1}{2}\begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3^n & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ -1 & 2 & -1 \\ -1 & 0 & 1 \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3^n & 0 & 3^n \\ -1 & 2 & -1 \\ -1 & 0 & 1 \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} 3^n+1 & 0 & 3^n-1 \\ 3^n-1 & 2 & 3^n-1 \\ 3^n-1 & 0 & 3^n+1 \end{bmatrix}
\end{aligned}$$

Question 5 **Properties of Eigenvalues**

For a given eigenvalue, the corresponding eigenvector may not be unique. Is it true that for every eigenvalue there is a unique *unit* eigenvector?

Solution. No. Consider the identity matrix $\mathbf{I} \in \mathbb{R}^{2\times 2}$. Every vector is an eigenvector, with eigenvalue 1, as $\mathbf{Ix} = \mathbf{x}$ hold for any $\mathbf{x} \in \mathbb{R}^2$. So any two distinct unit vectors provides a counterexample, like $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\end{bmatrix}$.

Question 6 **Singular Value Decomposition**

Compute the singular value decomposition of the matrix

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

Solution. We need to compute the eigenbasis of $\mathbf{A}^T\mathbf{A}$ to obtain $\mathbf{V}$, and the eigenbasis for $\mathbf{AA}^T$ to obtain $\mathbf{U}$. The matrix $\mathbf{\Sigma}$ is given by the square roots of the non-zero eigenvalues of either $\mathbf{A}^T\mathbf{A}$ or $\mathbf{AA}^T$, padded with zeros to be the same shape as $\mathbf{A}$.
So,

$$\mathbf{AA}^T = \begin{bmatrix}1 & 0 & 1\\0 & 1 & 1\\1 & 1 & 2\end{bmatrix}$$

We solve for the eigenvalues by constructing the characteristic equation,

$$\det(\mathbf{AA}^T - \lambda\mathbf{I}) = (1-\lambda)^2(2-\lambda) - 2(1-\lambda)$$

and solving for when $\det(\mathbf{AA}^T - \lambda\mathbf{I}) = 0$, we obtain the solutions

$$\lambda = 3, 1, 0$$

For $\lambda = 3$, we compute the eigenspace by solving $(\mathbf{AA}^T - 3\mathbf{I})\mathbf{x} = \mathbf{0}$.

$$\left[\begin{array}{ccc|c}1-3 & 0 & 1 & 0\\0 & 1-3 & 1 & 0\\1 & 1 & 2-3 & 0\end{array}\right] \sim \left[\begin{array}{ccc|c}1 & -1 & 0 & 0\\0 & 1 & \frac{-1}{2} & 0\\0 & 0 & 0 & 0\end{array}\right]$$

We obtain the solution set

$$E_3 = \text{span}[\begin{bmatrix}1\\1\\2\end{bmatrix}]$$

We need to choose a unit eigenvector from this set, so we choose

$$\mathbf{v}_3 = \frac{1}{\sqrt{6}}\begin{bmatrix}1\\1\\2\end{bmatrix}$$

For $\lambda = 1$, we compute the eigenspace by solving $(\mathbf{AA}^T - \mathbf{I})\mathbf{x} = \mathbf{0}$.

$$\left[\begin{array}{ccc|c}1-1 & 0 & 1 & 0\\0 & 1-1 & 1 & 0\\1 & 1 & 2-1 & 0\end{array}\right] \sim \left[\begin{array}{ccc|c}1 & 1 & 0 & 0\\0 & 0 & 1 & 0\\0 & 0 & 0 & 0\end{array}\right]$$

We obtain the solution set

$$E_1 = \text{span}[\begin{bmatrix}-1\\1\\0\end{bmatrix}]$$

and choose a unit eigenvector from this span,

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

For $\lambda = 0$, we compute the eigenspace by solving $\mathbf{A}\mathbf{A}^T\mathbf{x} = \mathbf{0}$.

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 2 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

We obtain the solution set

$$E_0 = \text{span}[\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}]$$

and choose a unit eigenvector from this span,

$$\mathbf{v}_0 = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

We have now obtained the eigenspectra of $\mathbf{A}$, which provides an orthonormal basis of $\mathbb{R}^3$, as $\mathbf{A}\mathbf{A}^T$ is symmetric, by the spectral theorem. We stack the three eigenvectors together into a matrix, and call the result $\mathbf{U}$.

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

We create a diagonal matrix out of the eigenvalues (in the same order as the eigenvectors) and call the result $\mathbf{D}$.

$$\mathbf{D} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

We thus have an eigenvalue decomposition of $\mathbf{A}\mathbf{A}^T$, by an orthogonal matrix $\mathbf{U}$ and diagonal matrix $\mathbf{D}$. (You can verify that $\mathbf{U}^{-1} = \mathbf{U}^T$).

$$\mathbf{A}\mathbf{A}^T = \mathbf{U}\mathbf{D}\mathbf{U}^T$$

Now, repeat the process for $\mathbf{A}^T\mathbf{A}$.

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

We solve for the eigenvalues by constructing the characteristic equation,

$$\det(\mathbf{A}^T\mathbf{A} - \lambda\mathbf{I}) = (2 - \lambda)^2 - 1$$

and solving for when $\det(\mathbf{A}^T\mathbf{A} - \lambda\mathbf{I}) = 0$, we obtain the solutions

$$\lambda = 3, 1$$

For $\lambda = 3$, we compute the eigenspace by solving $(\mathbf{A}^T\mathbf{A} - 3\mathbf{I})\mathbf{x} = \mathbf{0}$.

$$\left[\begin{array}{cc|c} 2-3 & 1 & 0 \\ 1 & 2-3 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} -1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

We obtain the solution set

$$E'_3 = \text{span}[\begin{bmatrix}1\\1\end{bmatrix}]$$

and choose a unit vector from this span,

$$\mathbf{v}'_3 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$$

For $\lambda = 1$, we compute the eigenspace by solving $(\mathbf{A}^T\mathbf{A} - \mathbf{I})\mathbf{x} = \mathbf{0}$.

$$\left[\begin{array}{cc|c}2-1 & 1 & 0\\1 & 2-1 & 0\end{array}\right] \sim \left[\begin{array}{cc|c}1 & 1 & 0\\0 & 0 & 0\end{array}\right]$$

We obtain the solution set

$$E'_1 = \text{span}[\begin{bmatrix}-1\\1\end{bmatrix}]$$

and choose a unit vector from this span,

$$\mathbf{v}'_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}-1\\1\end{bmatrix}$$

We have now obtained the eigenspectra of $\mathbf{A}$, which provides an orthonormal basis of $\mathbb{R}^2$, as $\mathbf{A}^T\mathbf{A}$ is symmetric, by the spectral theorem. We stack the two eigenvectors together, and call the result $\mathbf{V}$.

$$\mathbf{V} = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}$$

We stack the corresponding eigenvalues in the same order, and call the result $\mathbf{D}'$.

$$\mathbf{D}' = \begin{bmatrix}3 & 0\\0 & 1\end{bmatrix}$$

We thus have an eigenvalue decomposition of $\mathbf{A}^T\mathbf{A}$, by an orthogonal matrix $\mathbf{V}$ and diagonal matrix $\mathbf{D}'$. (You can verify that $\mathbf{P}^{-1} = \mathbf{P}^T$).

$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\mathbf{D}'\mathbf{V}^T$$

Now, to obtain the singular value decomposition, we let

$$\mathbf{\Sigma} = \begin{bmatrix}\sqrt{3} & 0\\0 & 1\\0 & 0\end{bmatrix}$$

(the non-zero singular values (square roots of eigenvalues) of $\mathbf{A}^T\mathbf{A}$ (or of $\mathbf{A}\mathbf{A}^T$), padded with zeros to match the shape of $\mathbf{A}$. Hence, we have the SVD

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \begin{bmatrix}\frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{3}}\\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{3}}\\ \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}}\end{bmatrix}\begin{bmatrix}\sqrt{3} & 0\\0 & 1\\0 & 0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix}^T$$

which can be verified to mutiply back out to

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

Appendix: Induction

The idea behind a proof by induction is usually employed to prove some statement to be true for all natural numbers. Suppose we have some statement $P(n)$ over the natural numbers, for example,

$$P(n) := \text{The sum of the first } n \text{ natural numbers is } \frac{n(n+1)}{2}.$$

We can verify $P(1), P(2), P(3)$ and find them all to be true, so we conjecture that $P(n)$ is true for all $n \in \mathbb{N}$. We cannot check infinitely many statements, so we can use induction. To do a proof by induction, we have to check two things.

1. The base case, usually $P(0)$ (or sometimes $P(1)$).
2. The step case. We assume $P(k)$ for some unknown k, and prove $P(k+1)$ still holds under this assumption.

Since we have $P(0)$ from the base case, the step case tells me that $P(1)$ must also be true. But if $P(1)$ holds, then $P(2)$ holds, and the next, and the next ...

Thus, we have $P(n)$ is true for all $n \in \mathbb{N}$. The assumption of $P(k)$ is called the *inductive hypothesis.*

Let's see an example by proving the above claim.

Proof by induction.
For the base case, prove that $P(0)$ holds, that is, that $\sum_{i=0}^{0} i = \frac{0(0+1)}{2}$. Both sides of the equation can be verified to be zero.

For the step case, we assume that

$$\sum_{i=0}^{k} i = \frac{k(k+1)}{2}$$

and try to prove that

$$\sum_{i=0}^{k+1} = \frac{(k+1)(k+2)}{2}$$

Proof,

$$\begin{aligned}
&\sum_{i=0}^{k+1} i \\
&= (k+1) + \sum_{i=0}^{k} i \\
&= (k+1) + \frac{k(k+1)}{2} \text{ (by the inductive hypothesis)} \\
&= \frac{2(k+1)}{2} + \frac{k(k+1)}{2} \\
&= \frac{2(k+1) + k(k+1)}{2} \\
&= \frac{(k+2)(k+1)}{2}
\end{aligned}$$

as required.