A Detailed Guide to Matrix Decompositions

A Step-by-Step Walkthrough

1 Eigen Decomposition: Understanding a Matrix's Core

Intuitive Description: Imagine a matrix as a geometric transformation (like a rotation, stretch, or shear). For any given matrix, there are special vectors called **eigenvectors** whose direction does not change when the transformation is applied. The matrix only scales them by a factor, which is the corresponding **eigenvalue**. The decomposition $A = PDP^T$ for a symmetric matrix breaks the transformation down into its fundamental components: its principle directions of action (the eigenvectors in P) and their scaling factors (the eigenvalues in the diagonal matrix D).

Problem: Find the complete eigen decomposition of the symmetric matrix A:

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

1.1 Step 1: Find the Eigenvalues (λ)

We solve the characteristic equation $\det(A - \lambda I) = 0$.

$$\begin{vmatrix} 2-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 2-\lambda \end{vmatrix} = 0$$

Expanding the determinant along the first row:

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

$$(2-\lambda)((2-\lambda)^2 - 1) + 1(-(2-\lambda)) = 0$$

Factor out the common term $(2-\lambda)$:

$$(2-\lambda)[(\lambda^2 - 4\lambda + 3) - 1] = 0$$

$$(2-\lambda)(\lambda^2 - 4\lambda + 2) = 0$$

This gives one eigenvalue $\lambda_1 = 2$. We find the other two using the quadratic formula for $\lambda^2 - 4\lambda + 2 = 0$:

$$\lambda = \frac{-(-4) \pm \sqrt{(-4)^2 - 4(1)(2)}}{2(1)} = \frac{4 \pm \sqrt{8}}{2} = 2 \pm \sqrt{2}$$

The eigenvalues are $\lambda_1 = 2$, $\lambda_2 = 2 + \sqrt{2}$, and $\lambda_3 = 2 - \sqrt{2}$.

1.2 Step 2: Find the Eigenvectors (v)

For each eigenvalue, we find a basis for the null space of $(A - \lambda I)$.

For $\lambda_1 = 2$:

$$(A - 2I)v = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies -y = 0 \text{ and } -x - z = 0$$

This means $y = 0$ and $x = -z$. A corresponding eigenvector is $v_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Normalizing gives $u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$.

For $\lambda_2 = 2 + \sqrt{2}$:

$$(A - (2 + \sqrt{2})I)v = \begin{pmatrix} -\sqrt{2} & -1 & 0 \\ -1 & -\sqrt{2} & -1 \\ 0 & -1 & -\sqrt{2} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

This gives $y = -\sqrt{2}x$ and $y = -\sqrt{2}z$, which implies $x = z$. An eigenvector is $v_2 = \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}$.

Normalizing gives $u_2 = \frac{1}{2} \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}$.

For $\lambda_3 = 2 - \sqrt{2}$:

$$(A - (2 - \sqrt{2})I)v = \begin{pmatrix} \sqrt{2} & -1 & 0 \\ -1 & \sqrt{2} & -1 \\ 0 & -1 & \sqrt{2} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

This gives $y = \sqrt{2}x$ and $y = \sqrt{2}z$, which implies $x = z$. An eigenvector is $v_3 = \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}$.

Normalizing gives $u_3 = \frac{1}{2} \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}$.

1.3 Step 3: Construct the Final Decomposition ($A = PDP^T$)

The matrix P contains the normalized eigenvectors, and D contains the corresponding eigenvalues.

$$P = \begin{pmatrix} 1/\sqrt{2} & 1/2 & 1/2 \\ 0 & -\sqrt{2}/2 & \sqrt{2}/2 \\ -1/\sqrt{2} & 1/2 & 1/2 \end{pmatrix}, \quad D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2+\sqrt{2} & 0 \\ 0 & 0 & 2-\sqrt{2} \end{pmatrix}$$

2 Singular Value Decomposition (SVD)

Intuitive Description: SVD is a powerful factorization that works for *any* matrix. It states that any linear transformation $B = U\Sigma V^T$ can be understood as a sequence of three fundamental operations:

1. **A rotation** (V^T): An orthonormal matrix that aligns the input space.
2. **A scaling** (Σ): A diagonal matrix that stretches or shrinks the space along the new axes. The scaling factors are the **singular values**.
3. **Another rotation** (U): An orthonormal matrix that aligns the result into the final output space.

Problem: Find the SVD of $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$.

2.1 Step 1: Find U and Σ from BB^T

We can start by analyzing the symmetric matrix BB^T. Its eigenvectors are the columns of U, and its eigenvalues are the squares of the singular values ($\lambda = \sigma^2$).

$$BB^T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

Eigenvalues of BB^T: $\det(BB^T - \lambda I) = 0$

$$\begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 1-\lambda & 0 \\ 1 & 0 & 1-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda)^2 - 1(1-\lambda) - 1(-(1-\lambda)) = 0$$

$$(1-\lambda)[(2-\lambda)(1-\lambda)-1-1] = (1-\lambda)[\lambda^2-3\lambda+2-2] = (1-\lambda)(\lambda^2-3\lambda) = \lambda(1-\lambda)(\lambda-3) = 0$$

The eigenvalues are $\lambda_1 = 3, \lambda_2 = 1, \lambda_3 = 0$. The singular values are their square roots: $\sigma_1 = \sqrt{3}, \sigma_2 = 1$. ($\sigma_3 = 0$ is trivial).

Eigenvectors of BB^T **(columns of U)**: For $\lambda_1 = 3$: $(BB^T - 3I)u = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -2 & 0 \\ 1 & 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$. This implies $x = 2y$ and $x = 2z$. An eigenvector is $\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$. Normalizing gives $u_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$.

For $\lambda_2 = 1$: $(BB^T - 1I)u = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$. This implies $x = 0$ and $y = -z$. An eigenvector is $\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$. Normalizing gives $u_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$.

For $\lambda_3 = 0$: $(BB^T - 0I)u = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$. This implies $y = -x$ and $z = -x$. An eigenvector is $\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$. Normalizing gives $u_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$.

So, $U = \begin{pmatrix} 2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & -1/\sqrt{3} \end{pmatrix}$.

2.2 Step 2: Find V using U

The columns of V (right singular vectors) can be found using the formula $v_i = \frac{1}{\sigma_i} B^T u_i$.

$$v_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{18}} \begin{pmatrix} 3 \\ 3 \end{pmatrix} = \frac{3}{3\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$v_2 = \frac{1}{1} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

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

2.3 Step 3: Assemble the Full SVD ($B = U\Sigma V^T$)

$$U = \begin{pmatrix} 2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & -1/\sqrt{3} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad V^T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

---

3 Low-Rank Approximation: Compressing Information

Intuitive Description: SVD is powerful because it organizes a matrix by its "importance." The largest singular value and its vectors capture the most significant part of the transformation.

By keeping only the top k singular values, we create a rank-k matrix that is the best possible approximation of the original, which is key for data compression and noise reduction.

Problem: Find the best rank-1 approximation of the matrix B.

3.1 Step 1: Isolate the Most Significant Component

We use the largest singular value $\sigma_1 = \sqrt{3}$ and its corresponding singular vectors:

$$u_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \quad v_1^T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix}$$

3.2 Step 2: Compute the Rank-1 Approximation (B_1)

The best rank-1 approximation is $B_1 = \sigma_1 u_1 v_1^T$.

$$B_1 = \sqrt{3} \cdot \left(\frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \right) \cdot \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix} \right)$$

$$B_1 = \frac{\sqrt{3}}{\sqrt{12}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix} = \frac{\sqrt{3}}{2\sqrt{3}} \begin{pmatrix} 2 & 2 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 & 2 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$$

This rank-1 matrix captures the dominant structure of the original matrix B.