CS246 Exam 2025 — Question 4: Dimensionality Reduction (12 points)

Setup

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

Part 1 (3 points) — Finding v_1 and σ_1 via the Power Method

The power method applied to $A^\top A$ converges to the eigenvector of the largest eigenvalue. We're told it converges to $x = (1/\sqrt{2})(1, 1)^\top$.

v_1: The columns of V are the eigenvectors of $A^\top A$, so the converged vector is directly the first right singular vector:

$$\boxed{v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}}$$

σ_1: Since eigenvalues of $A^\top A$ equal σ_i^2, we compute $\lambda_1 = v_1^\top (A^\top A) v_1$:

$$A^T A \cdot v_1 = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 7/\sqrt{2} \\ 7/\sqrt{2} \end{pmatrix}$$

$$\lambda_1 = v_1^T \cdot (A^T A \cdot v_1) = \frac{1}{\sqrt{2}} \cdot \frac{7}{\sqrt{2}} + \frac{1}{\sqrt{2}} \cdot \frac{7}{\sqrt{2}} = \frac{7}{2} + \frac{7}{2} = 7$$

$$\boxed{\sigma_1 = \sqrt{\lambda_1} = \sqrt{7}}$$

Verification: The characteristic equation $(5-\lambda)^2 - 4 = 0$ gives $\lambda = 7$ and $\lambda = 3$, confirming $\sigma_1 = \sqrt{7}$. ✓

Part 2 (3 points) — Prove that $u_i = Av_i / \sigma_i$

Start from the SVD definition: $A = U\Sigma V^\top$

Step 1: Right-multiply by V: $AV = U\Sigma V^\top V = U\Sigma$ (since $V^\top V = I$)

Step 2: Look column by column: the i-th column of AV is Av_i, and the i-th column of UΣ is $\sigma_i u_i$ (Σ is diagonal). Therefore:

$$Av_i = \sigma_i\, u_i$$

Step 3: Divide by σ_i (valid since $\sigma_i \neq 0$):

$$\boxed{u_i = \frac{Av_i}{\sigma_i}} \qquad \blacksquare$$

Why this matters: Once you have V and Σ from $A^\top A$, you get U for free — no need to separately solve the eigenproblem for $AA^\top$.

Part 3a (4 points) — Identify SVD vs. CUR

Decomposition A:

$$M \approx \begin{pmatrix} 0.24 & -0.95 \\ 0.24 & 0.27 \\ 0.94 & 0.17 \end{pmatrix} \begin{pmatrix} 8.24 & 0 \\ 0 & 4.33 \end{pmatrix} \begin{pmatrix} 0.03 & 0.66 & 0.74 & 0.12 \\ -0.22 & 0.39 & -0.20 & -0.87 \end{pmatrix}$$

Decomposition B:

$$M \approx \begin{pmatrix} 2 & 4 \\ 0 & 0 \\ 6 & 0 \end{pmatrix} \begin{pmatrix} 0.17 & 0 \\ -0.08 & 0.25 \end{pmatrix} \begin{pmatrix} 0 & 5 & 6 & 0 \\ 1 & 0 & 2 & 4 \end{pmatrix}$$

Answer: Decomposition A = **SVD**, Decomposition B = **CUR**

Reason 1: In A, the middle matrix is **diagonal with decreasing entries** (8.24 > 4.33), matching the Σ matrix in SVD. In B, the middle matrix has an off-diagonal entry (−0.08), characteristic of CUR's U matrix (pseudo-inverse of the intersection submatrix).

Reason 2: In B, the left factor contains actual **columns of M** (e.g., $(2,0,6)^\top$ and $(4,0,0)^\top$ are columns 3 and 4), and the right factor contains actual **rows of M**. CUR by definition uses real columns and rows. In A, the factors contain abstract non-integer values — orthonormal singular vectors typical of SVD.

Part 3b (2 points) — Which Has Smaller Error?

SVD has the smaller Frobenius norm error, guaranteed by the **Eckart–Young–Mirsky theorem**:

the truncated SVD is the optimal rank-k approximation.

Why CUR is still useful: CUR uses actual rows and columns, so when the original matrix is **sparse**, C and R are also sparse — leading to much more efficient storage. Additionally, CUR components are directly interpretable as real data points, unlike SVD's abstract singular vectors.