

Lab 12: Singular Value Decomposition **Solutions**

EECS 245, Spring 2026 at the University of Michigan

due for completion at 11:59PM Ann Arbor Time on Monday, June 22nd, 2026

Name: _____

username: _____

Each lab worksheet will contain several activities, some of which will involve writing code and others that will involve writing math on paper. To receive credit for a lab, you must complete as many of the activities as you can in 2 hours and submit a PDF of your work to Gradescope. We will provide specific instructions on how to submit programming activities (e.g. submitting the notebook or including a screenshot of some output).

Feel free to work with others in the course, but you must submit individually.

Recap: Singular Value Decomposition ([Chapters 10.1](#) and [10.2](#))

Suppose X is any $n \times d$ matrix. Then, there exists a **singular value decomposition (SVD)** of X of the form

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

where:

Matrix	Shape	Values Come From	Properties
U	$n \times n$	Columns are eigenvectors of XX^T , called the left singular vectors of X	Orthogonal, $U^T U = U U^T = I_{n \times n}$
Σ	$n \times d$	Each singular value σ_i is the square root of the i -th largest eigenvalue of $X^T X$ (or XX^T)	Diagonal, with value in position (i, i) equal to σ_i for $i = 1, 2, \dots, r = \text{rank}(X)$
V	$d \times d$	Columns are eigenvectors of $X^T X$, called the right singular vectors of X	Orthogonal, $V^T V = V V^T = I_{d \times d}$

If $\vec{u}_i$ and $\vec{v}_i$ are the i -th left and right singular vectors of X , respectively, then $X\vec{v}_i = \sigma_i\vec{u}_i$.

Activity 1: SVD Fundamentals

Suppose the matrix X has the singular value decomposition $X = U\Sigma V^T$ where

$$U = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} 1/\sqrt{2} & | & 0 \\ 0 & \vec{v}_2 & 1 \\ 1/\sqrt{2} & | & 0 \end{bmatrix}$$

a) How many rows and columns does X have? What is $\text{rank}(X)$?

Solution: X is a 2×3 matrix, matching the dimensions of Σ . $\text{rank}(X)$ is 2, since there are 2 non-zero singular values.

b) Find $\vec{v}_2$.

Solution: $\vec{v}_2$ is the eigenvector of $X^T X$ corresponding to the eigenvalue 2. It's also a unit vector, meaning we need to find a unit vector that's orthogonal to the other two columns, so $\vec{v}_2 = \begin{bmatrix} \sqrt{2}/2 \\ 0 \\ -\sqrt{2}/2 \end{bmatrix}$

c) Given that the first column of X and third column of X sum to $\begin{bmatrix} 0 \\ 5 \end{bmatrix}$, find σ_1 .

Solution:

$$X\vec{v}_1 = \sigma_1\vec{u}_1$$
$$X \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} = \sigma_1 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

The left side of the equation is the sum of X 's first and third columns, scaled by $1/\sqrt{2}$, so we can use the fact given in the problem here:

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 5 \end{bmatrix} = \sigma_1 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$\begin{bmatrix} 0 \\ 5/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 \\ \sigma_1 \end{bmatrix}$$
$$\sigma_1 = \frac{5\sqrt{2}}{2}$$

Activity 2: Outer Products

Consider the rank-2 matrix $X = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 3 & 3 \end{bmatrix}$.

- a) Write X as a sum of two rank-1 outer products, e.g. $X = \vec{x}_1 \vec{y}_1^T + \vec{x}_2 \vec{y}_2^T$.

Solution: X has a unique first column, while the second and third columns are the same. So, we can use a similar idea to the CR decomposition:

$$X = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}$$

- b) Find XX^T and $X^T X$, and the trace and determinant of each. Feel free to use numpy.

Solution:

$$XX^T = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 3 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 2 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 1+4+4 & 1+6+6 \\ 1+6+6 & 1+9+9 \end{bmatrix}$$

$$= \begin{bmatrix} 9 & 13 \\ 13 & 19 \end{bmatrix}$$

$$\det(XX^T) = 9 \cdot 19 - 13 \cdot 13 = 2$$

$$\text{trace}(XX^T) = 9 + 19 = 28$$

$$X^T X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ 1 & 3 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 1+1 & 2+3 & 2+3 \\ 2+3 & 4+9 & 4+9 \\ 2+3 & 4+9 & 4+9 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 5 & 5 \\ 5 & 13 & 13 \\ 5 & 13 & 13 \end{bmatrix}$$

$$\det(X^T X) = 0, X^T X \text{ is not invertible}$$

$$\text{trace}(X^T X) = 2 + 13 + 13 = 28$$

- c) If X is any $n \times d$ matrix, which of the following are guaranteed to be true, and why? *Hint: How*

does the trace of a matrix relate to its eigenvalues? How are the eigenvalues of XX^T and $X^T X$ related?

$$\text{trace}(XX^T) = \text{trace}(X^T X)$$

$$\det(XX^T) = \det(X^T X)$$

Solution: XX^T and $X^T X$ share the same **non-zero** values. Since the trace is equivalent to the sum of the eigenvalues, the non-zero eigenvalues being the same results in the same sum, as the zero eigenvalues don't affect it.

However, the same is not true for the determinant. In the case where $n \neq d$, one of XX^T or $X^T X$ will be larger. The larger matrix **must** have a zero eigenvalue because $\text{rank}(XX^T) = \text{rank}(X^T X)$, so its determinant will be 0. We can't guarantee that $X^T X$ won't be full rank, so its possible for the determinant to not be 0.

Activity 3: Rotating and Stretching

Suppose X is a 5×2 matrix with singular value decomposition $X = U\Sigma V^T$, and that $\vec{v}_1$ and $\vec{v}_2$ are the first and second columns of V , respectively. Furthermore, suppose $\vec{w} \in \mathbb{R}^2$ is a vector such that

$$\vec{w} = 3\vec{v}_1 - \vec{v}_2$$

a) Find $V^T \vec{w}$.

Solution:

$$\begin{aligned} V^T \vec{w} &= V^T (3\vec{v}_1 - \vec{v}_2) \\ &= 3V^T \vec{v}_1 - V^T \vec{v}_2 \\ &= 3 \begin{bmatrix} \vec{v}_1 \cdot \vec{v}_1 \\ \vec{v}_2 \cdot \vec{v}_1 \end{bmatrix} - \begin{bmatrix} \vec{v}_1 \cdot \vec{v}_2 \\ \vec{v}_2 \cdot \vec{v}_2 \end{bmatrix} \end{aligned}$$

We can simplify this further thanks to V being an orthogonal matrix. $\vec{v}_1$ and $\vec{v}_2$ are orthogonal, and their norms are both 1.

$$\begin{aligned} V^T \vec{w} &= 3 \begin{bmatrix} \|\vec{v}_1\|^2 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ \|\vec{v}_2\|^2 \end{bmatrix} \\ &= 3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ -1 \end{bmatrix} \end{aligned}$$

b) Suppose X 's two singular values are $\sigma_1 = 10$ and $\sigma_2 = 3$. Find $\Sigma V^T \vec{w}$.

Solution:

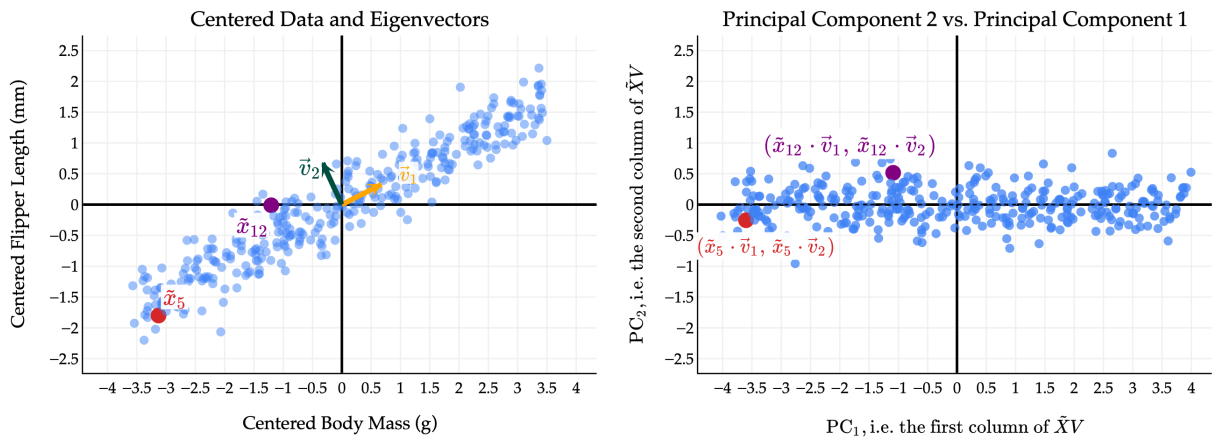
$$\begin{aligned} \Sigma V^T \vec{w} &= \Sigma \begin{bmatrix} 3 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 10 & 0 \\ 0 & 3 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 30 \\ -3 \\ 0 \\ 0 \\ 0 \end{bmatrix} \end{aligned}$$

c) Let $\vec{z} = \Sigma V^T \vec{w}$. In English, what does $\vec{z}$ represent, relative to $\vec{w}$?

Solution: $V^T \vec{w}$ rotates $\vec{w}$ using the right singular vectors as the basis. Since $\vec{w}$ is already composed of those vectors, it returns the coefficients back in a vector. Σ then shifts the resulting vector by scaling each component using the singular values.

Recap: Principal Component Analysis (Chapters 10.3 and 10.4)

- The goal of **principal component analysis (PCA)** is reducing the dimensionality of a dataset by constructing new features — called **principal components**.
- These new features are linear combinations of the existing features in the data, and are constructed to **minimize the mean squared orthogonal error** of the data when projected onto the new features.
- As we see in Chapter 10.3, this is equivalent to finding the **directions along which the data is most spread out**.



The plot on the left shows the direction vectors $\vec{v}_1$ and $\vec{v}_2$ which define the first and second principal components, respectively. $\vec{v}_1$ is the direction that captures the most variability, followed by $\vec{v}_2$. Note that $\vec{v}_1$ and $\vec{v}_2$ are **orthogonal**, which results in the principal components (new features) being **uncorrelated**, as we see in the plot on the right.

The “PCA recipe” is as follows:

1. Starting with an $n \times d$ matrix X of n data points in d dimensions and **mean-center** the data by subtracting the mean of each column from itself. The new matrix is $\tilde{X}$.
2. Compute the singular value decomposition of $\tilde{X}$: $\tilde{X} = U\Sigma V^T$.
3. **The columns of V (rows of V^T) describe the directions of maximal variance in the data!** For instance, the single “best direction” is the eigenvector of $\tilde{X}\tilde{X}^T$ with the largest eigenvalue, i.e. $\vec{v}_1$ in $\tilde{X} = U\Sigma V^T$.
4. Principal component (new feature) j comes from multiplying $\tilde{X}$ by the j -th column of V .

$$PC_j = \tilde{X}\vec{v}_j = \sigma_j\vec{u}_j$$

5. The variance of the new feature is

$$\text{Var}(PC_j) = \frac{\sigma_j^2}{n}$$

The proportion of total variance in $\tilde{X}$ that is explained by PC_j is

$$\text{proportion of variance explained by PC } j = \frac{\sigma_j^2}{\sum_{k=1}^r \sigma_k^2}$$

Activity 4: PCA Practice

Suppose A , B , and C are each 100×2 matrices, representing $n = 100$ points in $\mathbb{R}^2$. The three datasets are shown in the scatter plots below. (Matrix A is in Plot A, matrix B is in Plot B, and matrix C is in Plot C.)



Assume that A , B , and C are each already mean-centered.

- a) If we applied PCA to each of the above datasets, and created just one principal component in each case, for which dataset would the first principal component have the smallest mean squared orthogonal error — A , B , or C ?

Solution: B has a strong negative correlation, with little spread in the perpendicular direction, while A and C 's dominant directions aren't as clear.

- b) Suppose $\tilde{X} = U\Sigma V^T$ is the singular value decomposition of $\tilde{X}$, and that

$$\Sigma = \begin{bmatrix} 16 & 0 \\ 0 & 4 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}, \quad \underbrace{V = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}}_{\text{not } V^T}$$

Which dataset is most likely to be $\tilde{X}$ — A , B , or C ?

Solution: The singular values tell us that one direction is being scaled by 16, while the other is scaled by 4, meaning most of the variance is captured in the $\vec{v}_1$ direction. $\vec{v}_1$ has a positive x component and a negative y component, which matches the slope of plot B 's dominant direction, since A and C 's dominant directions are more positive.

- c) Again, suppose $\tilde{X} = U\Sigma V^T$ is the singular value decomposition of $\tilde{X}$, and that

$$\Sigma = \begin{bmatrix} 16 & 0 \\ 0 & 4 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}, \quad V = \underbrace{\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}}_{\text{not } V^T}$$

What is the proportion of the total variance in $\tilde{X}$ that is accounted for by the first principal component?

Solution:

$$\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} = \frac{16^2}{16^2 + 4^2} = \frac{256}{272}$$

- d) Suppose that in the graph of principal component 2 vs. principal component 1 (i.e. with PC 1 on the x -axis and PC 2 on the y -axis), a particular data point is plotted at $(4, 2)$. What is the corresponding point in the original (mean-centered) dataset? Your answer should be a tuple of two numbers, (x, y) (or equivalently, a vector in $\mathbb{R}^2$). *Hint: Start by understanding the plot on Page 4.*

Solution: There are two ways to approach this problem:

Option 1: Dot Products and Transformation

In the PC2 vs PC1 plot in the recap, you may have noticed the coordinates of $\vec{x}_{12}$ were $(\vec{x}_{12} \cdot \vec{v}_1, \vec{x}_{12} \cdot \vec{v}_2)$. This comes from multiplying $V^T \vec{x}_{12}$:

$$\begin{aligned} V^T \vec{x}_{12} &= \begin{bmatrix} - & \vec{v}_1 & - \\ - & \vec{v}_2 & - \end{bmatrix} \vec{x}_{12} \\ &= \begin{bmatrix} \vec{v}_1 \cdot \vec{x}_{12} \\ \vec{v}_2 \cdot \vec{x}_{12} \end{bmatrix} \end{aligned}$$

If this looks familiar, that's because it's similar to 3c! In 3c, $V^T \vec{w}$ rotated $\vec{w}$ by translating it into a space with $\{\vec{v}_1, \vec{v}_2\}$ as its basis. So, if we know a point's coordinates in $\vec{v}_1, \vec{v}_2$ space, we can work backwards to solve for the point's coordinates in x, y space:

$$\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} \vec{v}_1 \cdot \vec{w} \\ \vec{v}_2 \cdot \vec{w} \end{bmatrix}$$

$$\begin{aligned} 4 &= \vec{v}_1 \cdot \vec{w} = \frac{2}{\sqrt{5}}w_1 + \frac{-1}{\sqrt{5}}w_2 \\ 2 &= \vec{v}_2 \cdot \vec{w} = \frac{1}{\sqrt{5}}w_1 + \frac{2}{\sqrt{5}}w_2 \end{aligned}$$

Now we have a system of equations to solve for the coordinate of $\vec{w}$:

$$4 = \frac{2}{\sqrt{5}}w_1 + \frac{-1}{\sqrt{5}}w_2$$

$$2 = \frac{1}{\sqrt{5}}w_1 + \frac{2}{\sqrt{5}}w_2$$

$$8 = \frac{4}{\sqrt{5}}w_1 + \frac{-2}{\sqrt{5}}w_2$$

$$8 + 2 = \frac{4}{\sqrt{5}}w_1 + \frac{1}{\sqrt{5}}w_1 + \frac{-2}{\sqrt{5}}w_2 + \frac{2}{\sqrt{5}}w_2$$

$$10 = \frac{5}{\sqrt{5}}w_1$$

$$w_1 = \frac{10\sqrt{5}}{5} = 2\sqrt{5}$$

$$4 = \frac{2}{\sqrt{5}}(2\sqrt{5}) + \frac{-1}{\sqrt{5}}w_2$$

$$4 = 4 + \frac{-1}{\sqrt{5}}w_2$$

$$w_2 = 0$$

$$2 = \frac{1}{\sqrt{5}}(2\sqrt{5}) + \frac{2}{\sqrt{5}}w_2$$

$$2 = 2 + \frac{2}{\sqrt{5}}w_2$$

$$w_2 = 0 \quad \checkmark$$

$$\vec{w} = \begin{bmatrix} 2\sqrt{5} \\ 0 \end{bmatrix}$$

Option 2: Thinking in Terms of Linear Combinations

$(4, 2)$ are coefficients in the basis $\{\vec{v}_1, \vec{v}_2\}$. Since we know $\vec{v}_1$ and $\vec{v}_2$'s directions in x, y space, getting the point's coordinates in x, y space just involves computing the linear combination, using the coordinates in $\vec{v}_1, \vec{v}_2$ space as coefficients:

$$\begin{aligned} 4\vec{v}_1 + 2\vec{v}_2 &= 4 \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix} + 2 \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \\ &= \begin{bmatrix} 8/\sqrt{5} \\ -4/\sqrt{5} \end{bmatrix} + \begin{bmatrix} 2/\sqrt{5} \\ 4/\sqrt{5} \end{bmatrix} \\ &= \begin{bmatrix} 10/\sqrt{5} \\ 0 \end{bmatrix} = \begin{bmatrix} 2\sqrt{5} \\ 0 \end{bmatrix} \end{aligned}$$