Discussion 3

1. SVD

Suppose we have a matrix $A \in \mathbb{R}^{m \times n}$ with rank r. It turns out that its SVD has multiple forms, all of which can be useful depending on the problem we're working on.

We define the compact SVD as follows:

$$\underbrace{A}_{m\times n} = \underbrace{U_r}_{m\times r} \underbrace{\Sigma_r}_{r\times r} \underbrace{V_r^\top}_{r\times n}.$$

Here, $\Sigma_r \in \mathbb{R}^{r\times r}$ is a diagonal matrix containing non-zero singular values of A.

$$\Sigma_r = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix},$$

with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$.

Next, $U_r \in \mathbb{R}^{m\times r}$ is given by,

$$U_r = \begin{bmatrix} \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r \end{bmatrix},$$

where u_i is a left singular vector corresponding to non-zero singular value, σ_i, for $i = 1, 2, \ldots, r$. The columns of U_r are orthonormal and together they span the columnspace of A.

Finally, $V_r^\top \in \mathbb{R}^{r\times n}$ is given by,

$$V_r^\top = \begin{bmatrix} \vec{v}_1^\top \\ \vec{v}_2^\top \\ \vdots \\ \vec{v}_r^\top \end{bmatrix},$$

where $\vec{v}_j$ is a right singular vector corresponding to non-zero singular value, σ_j for $j = 1, 2, \ldots, r$. The rows of $V_r^\top$ are orthonormal and span the rowspace of A. Equivalently the columns of V_r span the column space of $A^\top$.

The matrix A can be expressed as,

$$A = \sigma_1 \vec{u}_1 \vec{v}_1^\top + \sigma_2 \vec{u}_2 \vec{v}_2^\top + \ldots + \sigma_r \vec{u}_r \vec{v}_r^\top.$$

This is called the dyadic SVD, since it's expressed as the sum of dyads (matrices of the form $uv^\top$). Assume now that $m \geq n$.

Another type of SVD which might be more familiar is the full SVD of A which is defined as follows:

$$\underbrace{A}_{m\times n} = \underbrace{U}_{m\times m} \underbrace{\Sigma}_{m\times n} \underbrace{V^\top}_{n\times n}.$$

Here, $\Sigma \in \mathbb{R}^{m\times n}$ has non-diagonal entries as zero. The diagonal entries of Σ contain the singular values and we can write Σ in terms of Σ_r as,

$$\Sigma = \left[\begin{array}{c|c} \Sigma_r & 0_{r\times(n-r)} \\ \hline 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{array}\right]$$

Next, $U \in \mathbb{R}^{m\times m}$ is an orthonormal matrix. U can be expressed in terms of U_r as,

$$U = \underbrace{[U_r}_{m\times r} \quad \underbrace{\vec{u}_{r+1} \quad \dots \quad \vec{u}_m]}_{m\times(m-r)}$$

The columns $\vec{u}_{r+1}, \vec{u}_{r+2}, \dots, \vec{u}_n$ are left singular vectors corresponding to singular value 0, and together span the nullspace of $A^\top$.

Finally, $V^\top$ is an orthonormal matrix and can be expressed in terms of $V_r^\top$ as,

$$V^\top = \begin{bmatrix} V_r^\top \\ \vec{v}_{r+1}^\top \\ \vdots \\ \vec{v}_n^\top \end{bmatrix} \begin{array}{l} \}\quad r\times n \\ \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} \quad (n-r)\times n \end{array}$$

The rows $\vec{v}_{r+1}^\top, \vec{v}_{r+2}^\top, \dots, \vec{v}_n^\top$ when transposed are the right singular vectors corresponding to singular value 0, and together they span the nullspace of A.

(a) For this problem assume that $m > n > r$. Label each of the following as True or False:

(a) $UU^\top = I$

(b) $U^\top U = I$

(c) $V^\top V = I$

(d) $VV^\top = I$

(e) $U_r^\top U_r = I$

(f) $U_r U_r^\top = I$

(g) $V_r V_r^\top = I$

(h) $V_r^\top V_r = I$

(b) Find the compact SVD of A, given that it has the following full SVD:

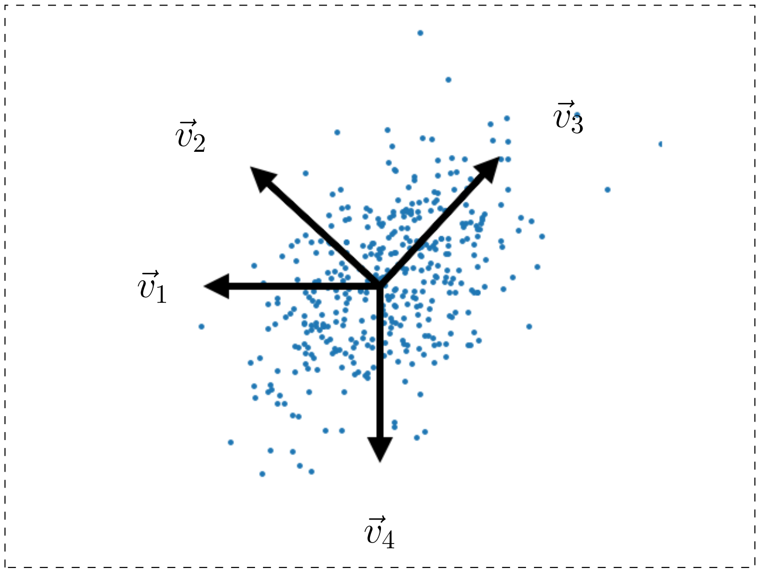
$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

(c) Find the full SVD of A, given that it has the following compact SVD:

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

2. PCA and Regression

(a) Given the following plot of data in $\mathbb{R}^2$ (i.e., each dot is a data point in $\mathbb{R}^2$) and candidate unit vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4 \in \mathbb{R}^2$, **identify the candidate vectors which could be the first principal component and second principal component of the data (and specify which is which).**



(b) Suppose we have data $\vec{x}_1, \ldots, \vec{x}_n \in \mathbb{R}^d$, where $n > d$. We arrange these data points into a matrix, i.e.,

$$X = \begin{bmatrix} \vec{x}_1^\top \\ \vdots \\ \vec{x}_n^\top \end{bmatrix} \in \mathbb{R}^{n \times d}. \tag{1}$$

Assume that X is centered, i.e., each column has mean zero: $(1/n)\sum_{i=1}^{n} \vec{x}_i = \vec{0}_d$, where $\vec{0}_d$ is the zero vector in $\mathbb{R}^d$. Suppose that X has compact SVD given by $X = U_d \Sigma_d V_d^\top$ where

$$U_d = \begin{bmatrix} \vec{u}_1, \ldots, \vec{u}_d \end{bmatrix} \in \mathbb{R}^{n \times d}, \qquad V_d = \begin{bmatrix} \vec{v}_1, \ldots, \vec{v}_d \end{bmatrix} \in \mathbb{R}^{d \times d}, \qquad \Sigma_d = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_d \end{bmatrix} \in \mathbb{R}^{d \times d} \tag{2}$$

where $\sigma_1 > \sigma_2 > \cdots > \sigma_d > 0$. **From this SVD, identify the top k principal components of the data $\{\vec{x}_1, \ldots, \vec{x}_n\} \subseteq \mathbb{R}^d$, where $k \leq d$.**

HINT: Recall that the first principal component solves the optimization problem $\underset{\vec{w} \in \mathbb{R}^d :\ \|\vec{w}\|_2 = 1}{\operatorname{argmax}} \quad \vec{w}^\top X^\top X \vec{w}$.

3. Singular Values

(a) Suppose $A \in \mathbb{R}^{3\times 2}$ is a matrix such that $A^\top A$ is given by

$$A^\top A = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}. \tag{3}$$

What are the singular values of A?

(b) Suppose that $B \in \mathbb{R}^{3\times 2}$ has singular values 0, $\sqrt{2}$, and $\sqrt{7}$. Let $C = \begin{bmatrix} B & -B & 3I_3 \end{bmatrix} \in \mathbb{R}^{3\times 7}$, where $I_3 \in \mathbb{R}^{3\times 3}$ is the 3×3 identity matrix. **What are the singular values of C?**

HINT: Consider the matrix $CC^\top \in \mathbb{R}^{3\times 3}$.