EECS 127/227AT Optimization Models in Engineering
Spring 2020

Discussion 1

1. Understanding ellipses

Consider the Euclidean space $\mathbb{R}^2$ with the orthogonal basis $\{\vec{e_1}, \vec{e_2}\}$. In this exercise, we study the ellipse

$$\mathcal{E} = \left\{ x_1\vec{e_1} + x_2\vec{e_2} \;\middle|\; x_1, x_2 \in \mathbb{R},\ \left(\sqrt{5}x_1 - \frac{3\sqrt{5}}{5}x_2\right)^2 + \left(\frac{4\sqrt{5}}{5}x_2\right)^2 \leq 8 \right\}.$$

1. Show that we can express the ellipse as $\mathcal{E} = \{\vec{x} \in \mathbb{R}^2 \mid \vec{x}^\top A\vec{x} \leq 1\}$ for symmetric positive definite A, where

$$A = \frac{1}{8}\begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

Solution:

$$\begin{aligned}
\mathcal{E} &= \{x_1\vec{e_1} + x_2\vec{e_2} \mid x_1, x_2 \in \mathbb{R},\ \left(\sqrt{5}x_1 - \frac{3\sqrt{5}}{5}x_2\right)^2 + \left(\frac{4\sqrt{5}}{5}x_2\right)^2 \leq 8\} \\
&= \{x_1\vec{e_1} + x_2\vec{e_2} \mid x_1, x_2 \in \mathbb{R},\ 5x_1^2 + 5x_2^2 - 6x_1x_2 \leq 8\} \\
&= \{\vec{x} \in \mathbb{R}^2 \mid \vec{x}^\top A\vec{x} \leq 1\},
\end{aligned}$$

for $A = \frac{1}{8}\begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix}$.

A is symmetric and has positive eigenvalues and is thus positive definite.

2. Show that the ellipse $\mathcal{E}$ can be viewed as a linear transformation of the unit disk by finding B such that $\mathcal{E} = \{B\vec{v} \mid \|\vec{v}\|_2 \leq 1\}$. Is this B unique?

Solution:

Since A is symmetric, by Spectral Theorem we can write $A = U\Lambda U^\top$ for orthogonal matrix U and diagonal matrix Λ. Observe that $C = \Lambda^{\frac{1}{2}}U^\top$ satisfies $A = C^\top C$ and is a square-root of A. Note that we can define $\Lambda^{1/2}$ since A is positive definite. But this is not the only square-root. For any orthogonal matrix V, $C = V\Lambda^{\frac{1}{2}}U^\top$ satisfies $C^\top C = A$. If we restrict to a symmetric square-root then we can take $C = U\Lambda^{\frac{1}{2}}U^\top$. Note that in all cases we can define $B = C^{-1} = U\Lambda^{-\frac{1}{2}}V^\top$.

We have,

$$\begin{aligned}
\mathcal{E} &= \{\vec{x} \in \mathbb{R},\ \vec{x}^\top A\vec{x} \leq 1\} \\
&= \{\vec{x} \in \mathbb{R}^2,\ \vec{x}^\top C^\top C\vec{x} \leq 1\} \\
&= \{\vec{x} \in \mathbb{R}^2,\ \|C\vec{x}\|_2^2 \leq 1\} \\
&= \{C^{-1}\vec{v},\ \|\vec{v}\|_2^2 \leq 1,\ \vec{v} \in \mathbb{R}^2\} \\
&= \{B\vec{v},\ \|\vec{v}\|_2 \leq 1,\ \vec{v} \in \mathbb{R}^2\}.
\end{aligned}$$

Another possible way to find B is to remark that C act as a change of variables on the original coordinates: $B = C^{-1}$ with $C = \frac{\sqrt{5}}{5}\begin{bmatrix} 5 & -3 \\ 0 & 4 \end{bmatrix}$.

3. Relate the length and direction of the semi-major and semi-minor axes of $\mathcal{E}$ to the singular values of B (or eigenvalues of A).

   Solution:
   We have $B = U\Lambda^{\frac{1}{2}}V^\top$. So we can view the effect of B on the unit disk as a rotation $V^\top$ followed by a scaling along the x_1 and x_2 axis (by diagonal entries of $\Lambda^{\frac{1}{2}}$ respectively), and followed by another rotation (U). The length of the semi axes of the ellipse is determined by the singular values of B and the direction/angle is determined by the left singular vectors (columns of U).
   Since $A = U\Lambda U^\top$, we can find corresponding relations to eigenvalues and eigenvectors of A.

4. Compute the area of $\mathcal{E}$.

   Solution: The area of $\mathcal{E}$ is:

$$\begin{aligned}
\text{Area}(\mathcal{E}) &= \text{Area}(\{B\vec{v} \in \mathbb{R}^2 \mid \|\vec{v}\|_2 \leq 1\}) \\
&= |\det(B)|\text{Area}(\{\vec{v} \in \mathbb{R}^2 \mid \|\vec{v}\|_2 \leq 1\}) \\
&= |\det(B)|\pi. \\
&= \frac{\pi}{2}
\end{aligned}$$

2. SVD

Suppose we have a matrix $A \in \mathbb{R}^{m\times n}$ with rank r.

We define the *compact (or "thin") SVD* of A 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 $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$.

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

$$U_r = \left[\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r\right],$$

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. Why? (**Exercise**).

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 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 $A^\top$. Why? (**Exercise**).

Note that the matrix A can be expressed as

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

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

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

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

$$U = \big[\underbrace{U_r}_{m\times r}\,\underbrace{\vec{u}_{r+1}\ldots\vec{u}_m}_{m\times(m-r)}\big].$$

The columns $\vec{u}_{r+1}, \vec{u}_{r+2}, \ldots, \vec{u}_n$ are left singular values corresponding to singular value 0 and together span the nullspace of $A^\top$. Why? (**Exercise**).
Finally, $V^\top$ is an orthogonal matrix and can be expressed in terms of $V_r^\top$ as,

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

The rows $\vec{v}_{r+1}^\top, \vec{v}_{r+2}^\top, \ldots, \vec{v}_n^\top$ when transposed are the right singular vectors corresponding to singular value of 0 and together span the nullspace of A. Why? (**Exercise**).

1. *Properties of the decomposition matrices.* Assume that $m > n > r$. Which of the following are true?
   (a) $UU^\top = I$

   Solution: True. U is an orthogonal matrix.

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

Solution: True. U is an orthogonal matrix.

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

Solution: True. V is an orthogonal matrix.

(d) $VV^\top = I$

Solution: True. V is an orthogonal matrix.

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

Solution: True. The columns of U_r are orthonormal.

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

Solution: False. $U_r U_r^\top$ is a $m \times m$, matrix but has rank less than or equal to r (since U_r has rank r and product of matrices has rank less than or equal to minimum of individual ranks). This is a projection matrix onto the range of A.

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

Solution: False. $V_r V_r^\top$ is a $n \times n$, matrix but has rank less than or equal to r (since V_r has rank r and product of matrices has rank less than or equal to minimum of individual ranks). This is a projection matrix onto the range of $A^\top$.

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

Solution: True. The rows of $V_r^\top$ are orthonormal.

When dealing with equations for matrices expressed in SVD form, it is crucial that you are clear on which of the previous matrix products simplify to the identity matrix and which do not.

2. *Building the compact SVD.* Let matrix A have full SVD

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

Find the compact SVD of A.

Solution: The thin SVD of A is given by:

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

3. *Building the full SVD.* Let matrix A have compact SVD

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

Find the full SVD of A.

Solution: Observe that in this case, the full SVD of A has Σ and $V^\top$ as those in the compact SVD but $U \in \mathbb{R}^{3\times 3}$. Thus we need to find a unit-norm column $\vec{u}_3$ orthogonal to columns of U_r. We can use a system of linear equations to solve this. That is we want $\vec{u}_3 = [x, \;\; y, \;\; z]$ so we must have

- $[1/\sqrt{2}, \;\; 1/\sqrt{2}, \;\; 0]^\top \vec{u}_3 = 0$
- $[0, \;\; 0, \;\; 1]^\top \vec{u}_3 = 0$
- $\|\vec{u}_3\|_2 = 1$

Check that $\vec{u}_3 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$ satisfies our requirements.

Thus the full SVD of A is given by:

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

Alternatively we can use Gram-Schmidt procedure to find $\vec{u}_3$. This has the added advantage of being useful when we want to find the full SVD when more than one singular vector is missing.