EECS 127/227AT Optimization Models in Engineering
Spring 2020 **Homework 2**

This homework is due Friday, February 7, 2020 at 23:00 (11pm).
Self grades are due Friday, February 14, 2020 at 23:00 (11pm).

This version was compiled on 2020-02-09 06:34.

Questions marked practice will not be graded.

Submission Format: Your homework submission should consist of a single PDF file that contains all of your answers (any handwritten answers should be scanned) as well as your IPython notebook with solutions saved as a PDF.

1. Proof of the Fundamental Theorem of Linear Algebra

In this question, we will prove the fundamental theorem of linear algebra. For any $A \in \mathbb{R}^{m\times n}$, let $\mathcal{N}(A)$, $\mathcal{R}(A)$ and $\operatorname{rank}(A)$ denote the null space, range and rank of A respectively.

For any subspace, $\mathcal{S}$ with dimension, $\dim(\mathcal{S})$, let $\mathcal{S}^\perp$ denote its the subspace orthogonal to $\mathcal{S}$.

The fundamental theorem of linear algebra states that,

$$\mathcal{N}(A) \oplus \mathcal{R}(A^\top) = \mathbb{R}^n.$$

The proof technique we employ will first show that,

$$\mathcal{N}(A) = (\mathcal{R}(A^\top))^\perp.$$

Then we will prove that we can find orthonormal vectors $\vec{e}_1, \vec{e}_2, \dots \vec{e}_n$ such that $\mathcal{N}(A) = \operatorname{span}(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_l)$ and $\mathcal{R}(A^\top) = \operatorname{span}(\vec{e}_{l+1}, \vec{e}_{l+2}, \dots, \vec{e}_n)$. As a corollary we get the rank-nullity theorem:

$$\dim(\mathcal{N}(A)) + \operatorname{rank}(A) = n.$$

(a) First, show that $\mathcal{N}(A) \subseteq (\mathcal{R}(A^\top))^\perp$. *Hint: Consider $\vec{u}$ in $\mathcal{N}(A), \vec{v} \in \mathcal{R}(A^\top)$ and show that $\vec{u}^\top \vec{v} = 0$.*

Solution: Take a vector $\vec{u}$ from the nullspace: $A\vec{u} = 0, \vec{u} \neq 0$, and a vector $\vec{w}$ from the rowspace: $A^\top \vec{v} = \vec{w}$. Now we can show orthogonality:

$$\langle \vec{w}, \vec{u} \rangle = \langle A^\top \vec{v}, \vec{u} \rangle = \vec{v}^\top A \vec{u} = \vec{v}^\top \vec{0} = 0$$

Therefore, $\vec{u} \in (\mathcal{R}(A^\top))^\perp$. Since this holds for every u in $\mathcal{N}(A)$ we have $\mathcal{N}(A) \subseteq (\mathcal{R}(A^\top))^\perp$. Note that this is not an equality since we only proved that every vector in the nullspace is orthogonal , not that these vectors are the **only** orthogonal ones.

(b) Now show that: $(\mathcal{R}(A^\top))^\perp \subseteq \mathcal{N}(A)$

Hint 1: Sometimes moving from symbols to words makes things clearer. Another way of stating what you want to prove is that any vector $\vec{v}$ that is orthogonal to all vectors in the range of $A^\top$, must satisfy $A\vec{v} = 0$.

Hint 2: Consider $\vec{v} \in (\mathcal{R}(A^\top))^\perp$. What can you say about $\vec{v}^\top A^\top$?

Solution: Consider $\vec{v} \in (\mathcal{R}(A^\top))^\perp$ and $\vec{u} \in \mathcal{R}(A^\top)$. They are related by $\langle \vec{v}, \vec{u} \rangle = \vec{v}^\top \vec{u} = 0$. Each column of $A^\top$ is in $\mathcal{R}(A^\top)$. So we have $\vec{v}^\top \vec{u} = 0$ if $\vec{u}$ is a column of $A^\top$.
This implies $\vec{v}^\top A^\top = 0$. Taking transposes gives $(A\vec{v})^\top = 0$ which implies that $\vec{v} \in \mathcal{N}(A)$. Since this holds for any $v \in (\mathcal{R}(A^\top))^\perp$ we have $(\mathcal{R}(A^\top))^\perp \subseteq \mathcal{N}(A)$.

(c) Let $\dim(\mathcal{N}(A)) = l$ and let $\vec{e}_1, \ldots, \vec{e}_l$ be an orthonormal basis for $\mathcal{N}(A)$. Consider an extension of the basis to an orthonormal basis, $\vec{e}_1, \ldots, \vec{e}_n$ for $\mathbb{R}^n$. We will prove that $\vec{e}_{l+1}, \ldots, \vec{e}_n$ form a basis for $\mathcal{R}(A^\top)$ and as a consequence, the dimension of $\mathcal{R}(A^\top)$ is $n - l$.

i. Show that $\mathcal{R}(A^\top)$ lies in the span of $\vec{e}_{l+1}, \ldots, \vec{e}_n$. To do this, first express any vector $\vec{u} \in \mathcal{R}(A^\top)$ in terms of the basis vectors e_i and use $\mathcal{N}(A) = (\mathcal{R}(A^\top))^\perp$, which you proved in parts (a) and (b).
Hint: If a vector $\vec{u}$ in a vector space is orthogonal to one of the basis vectors $\vec{e}_i$, what is the value of the coefficient α_i when writing $\vec{u} = \alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \ldots$?
Solution: We can obtain the coefficient attached to the basis vector $\vec{e}_i$ by finding the scalar projection (dot product) of $\vec{u}$ onto $\vec{e}_i$. Therefore, by projecting a vector onto an orthonormal one, we see that $\alpha_i = 0$.
Now, take any $\vec{u} \in \mathcal{R}(A^\top)$. We can express it as a linear combination of A's basis vectors:

$$\vec{u} = \sum_{i=1}^{n} \alpha_i \vec{e}_i.$$

But we also know from parts (a) and (b) that u is orthogonal to any vectors in the nullspace, which are spanned by/include the first ℓ basis vectors:

$$\vec{u}^\top \vec{e}_i = 0$$

for all $i \in \{1, 2, \ldots, l\}$. Therefore, $\vec{u}^\top \vec{e}_i = \alpha_i = 0$ for all $i \in \{1, 2, \ldots, l\}$ and subsequently

$$\vec{u} = \sum_{i=l+1}^{n} \alpha_i \vec{e}_i.$$

Therefore, any vector $\vec{u} \in \mathcal{R}(A^\top)$ can be spanned by $\vec{e}_{l+1} \ldots \vec{e}_n$, making $\mathcal{R}(A^\top)$ a subset of the span of $\vec{e}_{l+1} \ldots \vec{e}_n$.

ii. From part (i) we know that $\mathcal{R}(A^\top) \subseteq \operatorname{span}(\vec{e}_{l+1}, \ldots, \vec{e}_n)$, but we want something stronger. Show that in fact $\mathcal{R}(A^\top) = \operatorname{span}(\vec{e}_{l+1}, \ldots, \vec{e}_n)$.
Hint 1: First show that the dimension of $\mathcal{R}(A^\top)$ is the same as the dimension of the space spanned by the basis vectors $\vec{e}_{l+1}, \ldots, \vec{e}_n$, i.e., show $\dim(\mathcal{R}(A^\top)) = n - l$. You can show this via a contradiction: assume that $\dim(\mathcal{R}(A^\top)) = k < n - l$, and show that a vector $\vec{u} \notin \mathcal{R}(A^\top)$ and $\vec{u} \in Span\{\vec{e}_{l+1}, \ldots, \vec{e}_n\}$ cannot exist. For the proof by contradiction, one approach is to consider an orthonormal basis $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k$ for $\mathcal{R}(A^\top)$, so we can find non-zero $\vec{u}' = \vec{u} - \sum_{i=1}^{k} (\vec{f}_i^\top \vec{u}) \vec{f}_i$ that is orthogonal to $\mathcal{R}(A^\top)$. Does $\vec{u}'$ lie in $\mathcal{N}(A)$? Does $\vec{u}'$ also lie in $\operatorname{span}(\vec{e}_{l+1}, \ldots, \vec{e}_n)$? Does this lead to a contradiction?
Hint 2: Think of this in easily visualizable dimensions. Take $n - l = 3$ and $k = 2$.
Hint 3: You may use a fact that for two subspaces, $\mathcal{S}_1$ and $\mathcal{S}_2$, if $\mathcal{S}_1 \subseteq \mathcal{S}_2$ and $\dim(\mathcal{S}_1) = \dim(\mathcal{S}_2)$ then $\mathcal{S}_1 = \mathcal{S}_2$.
Solution: Assume the contrary and let $\vec{f}_1, \ldots, \vec{f}_k$ be an orthonormal basis for $\mathcal{R}(A^\top)$. Then, there exists $\vec{u}$ in the span of $\vec{e}_{l+1}, \ldots, \vec{e}_n$ such that $\vec{u} \notin \mathcal{R}(A^\top)$. From this, we get $\vec{u}' = \vec{u} - \sum_{i=1}^{k} (\vec{f}_i^\top \vec{u}) \vec{f}_i$ is non-zero and orthogonal to $\mathcal{R}(A^\top)$.

Visually, you can think of the range of $A^\top$ as the x, y plane, and $\vec{u}$ extends into 3D space outside this plane. From there, we can define a $\vec{u}'$ that is perpendicular to the x, y plane simply by using Gram-Shmidt on the basis vectors, for example.
Thus, $\vec{u}' \in (\mathcal{R}(A^\top))^\perp = \mathcal{N}(A)$. However, we also have $\vec{u}' \in \text{Span}(\{\vec{e}_i\})_{i=l+1}^n$ which is a contradiction as $\mathcal{N}(A)$ and $\text{Span}(\{\vec{e}_i\})_{i=l+1}^n$ are orthogonal to each other. Therefore, the dimension of $\mathcal{R}(A^\top)$ is at least $n-l$ and is exactly $n-l$ as it is contained in an $n-l$ dimensional space.
Since, we have $\mathcal{R}(A^\top) \subseteq \text{Span}(\vec{e}_{l+1} \dots \vec{e}_n)$, and $\dim(\mathcal{R}(A^\top)) = \dim(\vec{e}_{l+1} \dots \vec{e}_n) = n-l$, we can conclude that $\mathcal{R}(A^\top) = \text{Span}(\{\vec{e}_i\}_{i=1+1}^n)$.

(d) Using part (c) argue why $\mathcal{N}(A) \oplus \mathcal{R}(A^\top) = \mathbb{R}^n$ and why the rank nullity theorem holds.
Solution: We have found a basis $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ such that the first l vectors $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_l$ form a basis for $\mathcal{N}(A)$ and the last $n-l$ vectors form a basis for $\mathcal{R}(A^\top)$. Thus the claim $\mathcal{N}(A) \oplus \mathcal{R}(A^\top) = \mathbb{R}^n$ follows. Also,

$$\begin{aligned}
\dim(\mathcal{N}(A)) + \text{rank}(A) &= \dim(\mathcal{N}(A)) + \text{rank}(A^\top) \\
&= \dim(\mathcal{N}(A)) + \dim(\mathcal{R}(A^\top)) \\
&= l + n - l = n.
\end{aligned}$$

2. Eigenvectors of a symmetric matrix

Let $\vec{p}, \vec{q} \in \mathbb{R}^n$ be two linearly independent vectors, with unit norm ($\|\vec{p}\|_2 = \|\vec{q}\|_2 = 1$). Define the symmetric matrix $A \doteq \vec{p}\vec{q}^\top + \vec{q}\vec{p}^\top$. In your derivations, it may be useful to use the notation $c \doteq \vec{p}^\top \vec{q}$.

(a) Show that $\vec{p}+\vec{q}$ and $\vec{p}-\vec{q}$ are eigenvectors of A, and determine the corresponding eigenvalues.
Solution: We have

$$A\vec{p} = c\vec{p} + \vec{q}, \quad A\vec{q} = \vec{p} + c\vec{q},$$

from which we obtain

$$A(\vec{p} - \vec{q}) = (c-1)(\vec{p} - \vec{q}), \quad A(\vec{p} + \vec{q}) = (c+1)(\vec{p} + \vec{q}).$$

Thus $\vec{u}_\pm := \vec{p} \pm \vec{q}$ is an (un-normalized) eigenvector of A, with eigenvalue $c \pm 1$.

(b) Determine the nullspace and rank of A.
Solution: If $\vec{x} \in \mathbb{R}^n$ is in the nullspace of A we must have: $A\vec{x} = 0$.

$$0 = A\vec{x} = \vec{p}(\vec{q}^\top \vec{x}) + \vec{q}(\vec{p}^\top \vec{x}).$$

Since $(\vec{q}^\top \vec{x})$ and $(\vec{p}^\top \vec{x})$ are scalars we can rewrite this as:

$$0 = A\vec{x} = (\vec{q}^\top \vec{x})\vec{p} + (\vec{p}^\top \vec{x})\vec{q} = 0.$$

However, since $\vec{p}, \vec{q}$ are linearly independent, the fact that a linear combination of $\vec{p}, \vec{q}$ is zero implies that $\vec{p}^\top \vec{x} = \vec{q}^\top \vec{x} = 0$. Hence, the nullspace of A is the set of vectors orthogonal to $\vec{p}$ and $\vec{q}$.
Finally, we have from the fundamental theorem of linear algebra and the fact that A is symmetric, $\mathcal{R}(A) = \mathcal{R}(A^\top) = \mathcal{N}(A)^\perp$. From the previous part, we get that $\mathcal{N}(A) = \text{span}(p, q)^\perp$ which implies that $\mathcal{R}(A) = \text{span}(p, q)$. and since p and q are independent $\text{rank}(A) = 2$.

(c) Find an eigenvalue decomposition of A, in terms of $\vec{p}, \vec{q}$. *Hint:* use the previous two parts.
Solution: Since the rank is 2, we need to find a total of two non-zero eigenvalues.

First, we check that that $\lambda = c \pm 1$ is not 0. We have $\vec{p} - \vec{q} \neq 0$ which implies $\|\vec{p} - \vec{q}\|_2^2 > 0$ which means $\|\vec{p}\|_2^2 + \|\vec{q}\|_2^2 - 2\vec{p}^\top \vec{q} > 0$. Therefore, we have $c < 1$ and through a similar proof with $\vec{p} - \vec{q}$, we have $-c < 1$. From these two facts, we get $|c| < 1$.
Thus, we have found two linearly independent eigenvectors $\vec{u}_\pm = \vec{p} \pm \vec{q}$ that do not belong to the nullspace. Then, the eigenvalue decomposition is

$$A = (c-1)\vec{v}_-\vec{v}_-^\top + (c+1)\vec{v}_+\vec{v}_+^\top,$$

where $\vec{v}_\pm$ are the normalized vectors $\vec{v}_\pm = \vec{u}_\pm / \|\vec{u}_\pm\|_2$. Since

$$\|\vec{p} \pm \vec{q}\|_2^2 = \vec{p}^\top \vec{p} \pm 2\vec{p}^\top \vec{q} + \vec{q}^\top \vec{q} = 2(1 \pm c),$$

we have

$$\vec{v}_\pm = \frac{1}{\sqrt{2(1 \pm c)}}(\vec{p} \pm \vec{q}),$$

so that the eigenvalue decomposition becomes

$$A = \frac{1}{2}\left((\vec{p} + \vec{q})(\vec{p} + \vec{q})^\top - (\vec{p} - \vec{q})(\vec{p} - \vec{q})^\top\right).$$

(d) (Practice) Now consider general $\vec{p}, \vec{q}$ that are not necessarily norm 1. Write A as a function $\vec{p}, \vec{q}$ and their norms and the new eigenvalues as a function of $\vec{p}, \vec{q}$ and their norms.
Solution: We can scale the matrix: with $\vec{p}_n = \vec{p}/\|\vec{p}\|_2$, $\vec{q}_n = \vec{q}/\|\vec{q}\|_2$, and $A_n = \vec{p}_n \vec{q}_n^\top + \vec{q}_n \vec{p}_n^\top$, we have

$$A = \|\vec{p}\|_2 \|\vec{q}\|_2 A_n.$$

Since A is just a scaled version of A_n, whose eigenvalues we determined in previous parts, the eigenvalues of A are scaled accordingly: with $c = \bar{\vec{p}}^\top \bar{\vec{q}}$ the eigenvalues of A are given by,

$$\lambda_\pm = \|\vec{p}\|_2 \|\vec{q}\|_2 (c \pm 1) = \vec{p}^\top \vec{q} \pm \|\vec{p}\|_2 \|\vec{q}\|_2.$$

The unit norm eigenvectors of A are same as that of A_n which gives,

$$A = \|\vec{p}\|_2 \|\vec{q}\|_2 A_n = \frac{\|\vec{p}\|_2 \|\vec{q}\|_2}{2}\left((\bar{\vec{p}} + \bar{\vec{q}})(\bar{\vec{p}} + \bar{\vec{q}})^\top - (\bar{\vec{p}} - \bar{\vec{q}})(\bar{\vec{p}} - \bar{\vec{q}})^\top\right).$$

3. Norms

(a) Show that the following inequalities hold for any vector $\vec{x} \in \mathbb{R}^n$:

$$\frac{1}{\sqrt{n}} \|\vec{x}\|_2 \leq \|\vec{x}\|_\infty \leq \|\vec{x}\|_2 \leq \|\vec{x}\|_1 \leq \sqrt{n} \|\vec{x}\|_2 \leq n \|\vec{x}\|_\infty.$$

As an aside: note that we can interpret different norms as different ways of computing distance between two points $\vec{x}, \vec{y} \in \mathbb{R}^2$. The ℓ_2 norm is the distance as the crow flies (i.e. point-to-point distance), the ℓ_1 norm, also known as the Manhattan distance is the distance you would have to cover if you were to navigate from $\vec{x}$ to $\vec{y}$ via a rectangular street grid, and the ℓ_∞ norm is the maximum distance that you have to travel in either the north-south or the east-west direction.

Solution: We have

$$\|\vec{x}\|_2^2 = \sum_{i=1}^n x_i^2 \leq n \cdot \max_i \, x_i^2 = n \cdot \|\vec{x}\|_\infty^2.$$

Also, $\|\vec{x}\|_\infty \leq \sqrt{x_1^2 + \ldots + x_n^2} = \|\vec{x}\|_2$.

The inequality $\|\vec{x}\|_2 \leq \|\vec{x}\|_1$ is obtained after squaring both sides, and checking that

$$\sum_{i=1}^n x_i^2 \leq \sum_{i=1}^n x_i^2 + \sum_{i \neq j} |x_i x_j| = \left(\sum_{i=1}^n |x_i|\right)^2 = \|\vec{x}\|_1^2.$$

The condition $\|\vec{x}\|_1 \leq \sqrt{n}\|\vec{x}\|_2$ is due to the Cauchy-Schwarz inequality

$$|\vec{z}^\top \vec{y}| \leq \|\vec{y}\|_2 \cdot \|\vec{z}\|_2,$$

applied to the two vectors $y = (1, \ldots, 1)$ and $\vec{z} = |\vec{x}| = (|x_1|, \ldots, |x_n|)$.

Finally, $\sqrt{n}\|\vec{x}\|_2 \leq n\|\vec{x}\|_\infty$, is achieved by an algebraic manipulation of the first derived bound using the fact that $\sqrt{n} = \frac{n}{\sqrt{n}}$.

(b) Show that for any non-zero vector x,

$$\text{card}(\vec{x}) \geq \frac{\|\vec{x}\|_1^2}{\|\vec{x}\|_2^2},$$

where $\text{card}(\vec{x})$ is the *cardinality* of the vector $\vec{x}$, defined as the number of non-zero elements in $\vec{x}$. Find all vectors $\vec{x}$ for which the lower bound is attained.

Solution: Let us apply the Cauchy-Schwarz inequality with $\vec{z} = |\vec{x}|$ again, and with $\vec{y}$ a vector with $y_i = 1$ if $x_i \neq 0$, and $y_i = 0$ otherwise. We have $\|\vec{y}\|_2 = \sqrt{k}$, with $k = \text{card}(\vec{x})$. Hence

$$|\vec{z}^\top \vec{y}| = \|\vec{x}\|_1 \leq \|\vec{y}\|_2 \cdot \|\vec{z}\|_2 = \sqrt{k} \cdot \|\vec{x}\|_2,$$

which proves the result. The bound is attained for vectors with k non-zero elements, all with the same magnitude.

4. Distinct Eigenvalues, Orthogonal Eigenspaces

Let $A \in \mathbb{S}^n$ (i.e. the set of $n \times n$ symmetric matrices) and $(\lambda_1, \vec{u}_1), (\lambda_2, \vec{u}_2), \lambda_1 \neq \lambda_2$ be distinct eigen-pairs of A. Show that $\langle \vec{u}_1, \vec{u}_2 \rangle = 0$, i.e eigenspaces corresponding to distinct eigenvalues are mutually orthogonal.

Solution: It is useful to note the following equality:

$$\langle A\vec{x}, \vec{y} \rangle = \vec{x}^\top A^\top \vec{y} = \langle \vec{x}, A^\top \vec{y} \rangle. \tag{1}$$

Now comes the massaging of equations:

$$\begin{aligned}
\lambda_1 \langle \vec{u}_1, \vec{u}_2 \rangle &= \langle \lambda_1 \vec{u}_1, \vec{u}_2 \rangle && \text{Linearity of inner product} \\
&= \langle A\vec{u}_1, \vec{u}_2 \rangle && A\vec{u}_1 = \lambda_1 \vec{u}_1 \\
&= \langle \vec{u}_1, A^\top \vec{u}_2 \rangle && \text{Equation 1} \\
&= \langle \vec{u}_1, A\vec{u}_2 \rangle && A \in \mathbb{S}^n \\
&= \langle \vec{u}_1, \lambda_2 \vec{u}_2 \rangle && A\vec{u}_2 = \lambda_2 \vec{u}_2 \\
&= \lambda_2 \langle \vec{u}_1, \vec{u}_2 \rangle. && \text{Linearity of inner product}
\end{aligned}$$

$$\begin{aligned}
\lambda_1 \langle \vec{u}_1, \vec{u}_2 \rangle &= \lambda_2 \langle \vec{u}_1, \vec{u}_2 \rangle \\
\underbrace{(\lambda_1 - \lambda_2)}_{\neq 0} \langle \vec{u}_1, \vec{u}_2 \rangle &= 0.
\end{aligned}$$

Thus, $\langle \vec{u}_1, \vec{u}_2 \rangle = 0$ for any $\vec{u}_1, \vec{u}_2$ corresponding to different eigenvalues. Stated differently, unique eigenvalues correspond to orthogonal eigenvectors.

This, in combination with the fact that the geometric multiplicity and algebraic multiplicity of a symmetric matrix are equal, allows us to construct an orthonormal set of eigenvectors. First, find all the distinct eigenvalues and their respective eigenvectors. Then, for all eigenvalues with algebraic multiplicity > 1, we know that the respective eigenspace is spanned by k linearly independent eigenvectors. Utilizing Gram-Schmidt, we can construct an orthonormal set of eigenvectors from this basis for this eigenspace. Putting the eigenvectors from these two cases together, we have constructed the U matrix of the decomposition.

5. **PSD Matrices** In this problem, we will analyze properties of PSD matrices. Assume $A \in \mathbb{R}^{n \times n}$ is a symmetric matrix.

   (a) Show that $\forall \vec{x} \in \mathbb{R}^n, \vec{x}^\top A \vec{x} \geq 0 \iff$ all eigenvalues of A are non-negative.

   (b) Show that A having non-negative eigenvalues allows us to decompose $A = P^\top P$ where $P \in \mathbb{S}^n_+$ (i.e. the set of $n \times n$ positive semidefinite matrices).

   (c) (Practice) Show that any matrix of the form $B = C^\top C \succeq 0$.

   (d) If $A \succeq 0$, all diagonal entries of A are non-negative, $A_{ii} \geq 0$.

   Solution:

   (a) $\implies$:

      i. Solution 1: We can plug in the Spectral Decomposition here:

      $$\vec{x}^\top A \vec{x} = \vec{x}^\top U \Sigma U^\top \vec{x} = \vec{v}^\top \Sigma \vec{v} \geq 0,$$

      where $\vec{v} := U^\top \vec{x}$ is a rotated version of $\vec{x}$ since U is orthonormal. Now, we just need to convert that final quadratic into any eigenvalue of A, and we can do that by choosing a $\vec{v}$ that pulls out whichever eigenvalue we want (e.g. if we want the first eigenvalue, we can choose the first unit vector). To be thorough, we can then realize that the set of $\vec{x}$'s such that $U^\top \vec{x} = \vec{e}_i$ for any unit vector, will pull out the ith eigenvalue, thus satisfying definition 2.

ii. Solution 2: We can just use the definition of an eigenvalue:

$$\vec{x}^\top A\vec{x} = \vec{x}\lambda\vec{x} = \lambda\vec{x}^\top\vec{x} = \lambda\|\vec{x}\|_2^2$$

Since norms/anything squared is always non-negative, in order for $\lambda\|\vec{x}\|_2^2 \geq 0, \lambda$ must be non-negative.

$\Longleftarrow$: Using the Spectral Decomposition again, we arrive at the equation $\vec{v}^\top\Sigma\vec{v}$, which we can expand further:

$$\vec{v}^\top\Sigma\vec{v} = \sum_i \lambda_i v_i^2 \geq 0,$$

where the last inequality came from the fact that anything squared is non-negative and all eigenvalues are non-negative by assumption of the problem.

(b) With all non-negative eigenvalues, we are able to define a matrix $A^{\frac{1}{2}} = U\Sigma^{\frac{1}{2}}U^\top$, where $\Sigma^{\frac{1}{2}}$ is a diagonal matrix with the square roots of A's eigenvalues. Note that $A^{\frac{1}{2}}$ is PSD since its eigenvalues are still non-negative. Thus, with $P = A^{\frac{1}{2}}$, we can show the following:

$$P^\top P = (A^{\frac{1}{2}})^\top A^{\frac{1}{2}} = (U\Sigma^{\frac{1}{2}}U^\top)^\top U\Sigma^{\frac{1}{2}}U^\top = U\Sigma^{\frac{1}{2}}U^\top U\Sigma^{\frac{1}{2}}U^\top = U\Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}}U^\top = U\Sigma U^\top = A.$$

(c) We can plug in $B = C^\top C$ into the quadratic form as follows:

$$\vec{x}^\top A\vec{x} = \vec{x}^\top C^\top C\vec{x} = \langle C\vec{x}, C\vec{x}\rangle = \|C\vec{x}\|_2^2 \geq 0.$$

(d) The quadratic form $\vec{x}^\top A\vec{x} \geq 0$ applies for all vectors $\vec{x}$. Therefore, let's choose a vector that will pull out A_{ii}: the ith unit vector. $A\vec{e}_i$ pulls out the ith column $\vec{a}_i$, followed by $\vec{e}_i^\top\vec{a}_i$, which will pull out the ith element of the ith column. Therefore, $\vec{e}_i^\top A\vec{e}_i = A_{ii} \geq 0$.

6. SVD Transformation

In this problem we will interpret the linear map corresponding to a matrix $A \in \mathbb{R}^{n\times n}$ by looking at its singular value decomposition, $A = UDV^\top$. Recall that here $U, D, V \in \mathbb{R}^{n\times n}$ and U, V are orthonormal(orthogonal) matrices while D is a diagonal matrix. We will first look at how $V^\top, D$ and U each separately transform the unit circle $C = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ and then look at their effect as a whole. This problem has an associated jupyter notebook, "svd_transformation.ipynb" that contains several parts (b,c,d,e) of the problem. These sub-parts can be answered in the notebook itself in the space provided and can be submitted as a pdf using the 'Download as pdf' feature that jupyter notebook supports.

(a) Show that $V^\top\vec{x}$ represents $\vec{x}$ in the basis defined by the columns of V. Recall: $V^\top V = I$.

For rest of the problem we restrict ourselves to the case where $A \in \mathbb{R}^{2\times 2}$ and move to the Jupyter notebook.

Solution:

(a) Since the columns of V form a basis for $\mathbb{R}^n$, we can represent $\vec{x} \in \mathbb{R}^n$ as $V\vec{z}$ for some $\vec{z}$ in $\mathbb{R}^n$. Then,

$$\begin{aligned} V^\top\vec{x} &= V^\top V\vec{z} \\ &= \vec{z}. \end{aligned}$$

The last equality follows since V is an orthogonal matrix.

The solutions for rest of the parts can be found in the Jupyter notebook solution.

7. Homework process

Whom did you work with on this homework? List the names and SIDs of your group members.