Numerics 1: HW 3

Cooper Simpson

September 17, 2020

Problem 1:

Prove that if $\mathbf{A}$ is positive definite, then $\mathbf{A}$ is non-singular.

Answer:

Proof.
We know that $\mathbf{A} \in \mathcal{C}^{n\times n}$ is positive definite if and only if the following expression holds.

$$\mathbf{x}^T\mathbf{A}\mathbf{x} > 0,\ \forall \mathbf{x} \in \mathcal{C}^n,\ \mathbf{x} \neq 0$$

We want to show that $\mathbf{A}$ is then non-singular. We will do this by proving the contrapositive of our statement – $\mathbf{A}$ singular implies $\mathbf{A}$ not positive definite.

We assume that $\mathbf{A} \in \mathcal{C}^{n\times n}$ is singular.

$$\implies \exists \mathbf{x} \in \mathcal{C}^n,\ \mathbf{x} \neq 0,\ s.t.\ \mathbf{A}\mathbf{x} = \mathbf{0}$$

This holds because the two statements are equivalent as stated in class.

Let $\mathbf{x}_1 \in \mathcal{C}^n$ be given such that $\mathbf{A}\mathbf{x}_1 = \mathbf{0}$ and $\mathbf{x}_1 \neq \mathbf{0}$.

$$\implies \mathbf{x}_1^T\mathbf{A}\mathbf{x}_1 = \mathbf{x}_1^T\mathbf{0} = 0$$

We have found a at least one vector such that the inner product $\langle \mathbf{A}\mathbf{x}, \mathbf{x}\rangle$ is not positive. Also, because the definition for positive definite is bidirectional, we can use this in the reverse direction...

$\therefore$ $\mathbf{A}$ is not positive definite.

Having proved the contrapositve this implies that if $\mathbf{A}$ is positive definite, then $\mathbf{A}$ is non-singular. $\square$

Problem 2:

Prove that for $\mathbf{A} = \mathbf{M}^T\mathbf{M}$, where $\mathbf{M}$ is any real square $(n \times n)$ non-singular matrix, $\mathbf{A}$ is positive definite.

Answer:

Proof.
Let $\mathbf{A} = \mathbf{M}^T\mathbf{M}$, where $\mathbf{M}$ is a real $n \times n$ non-singular matrix.

We want to show that $\mathbf{A}$ is positive definite. To do this we need to show the following:

$$\forall \mathbf{x} \in \mathcal{R}^n,\ \mathbf{x} \neq \mathbf{0},\ \mathbf{x}^T\mathbf{A}\mathbf{x} > 0$$

We note that we are considering $\mathcal{R}^n$ because $\mathbf{M}$ is real and therefore so is $\mathbf{A}$.

Let $\mathbf{x} \in \mathcal{R}^n$ be given such that $\mathbf{x} \neq \mathbf{0}$.

$\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{M}^T\mathbf{M}\mathbf{x}$

$= (\mathbf{M}\mathbf{x})^T\mathbf{M}\mathbf{x} = \langle \mathbf{M}\mathbf{x}, \mathbf{M}\mathbf{x} \rangle$

Where the last expression above is the inner product.

Properties of the inner product $\implies \langle \mathbf{M}\mathbf{x}, \mathbf{M}\mathbf{x} \rangle > 0$

$\implies \mathbf{x}^T\mathbf{A}\mathbf{x} > 0$

$\therefore$ the matrix $\mathbf{A}$ is psitive definite. $\square$

Problem 3:

Let $\mathbf{A}$ be a real $n \times m$ matrix with rank r.

a:

Write down and describe the singular value decomposition for $\mathbf{A}$.

Answer:
The singular value decomposition of a rank r matrix $\mathbf{A} \in \mathcal{R}^{n \times n}$ is given as follows:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

$\mathbf{U}$ is an $n \times n$ orthogonal matrix, $\mathbf{V}^T$ is an $m \times m$ orthogonal matrix, and $\mathbf{\Sigma}$ is an $n \times m$ rectangular diagonal matrix. We note that an orthogonal matrix is a matrix with orthonormal vectors. Also, we are using the term rectangular diagonal matrix to indicate that the only elements that may be non-zero are of the form $\mathbf{\Sigma}_{i,i}$ – on the diagonal starting from the top left. Furthermore, these values on the "diagonal" of $\mathbf{\Sigma}$ are called the singular values and have the following property:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

In other words the first r (as in the rank) values on the diagonal are positive and decreasing. We note that $r \leq min\{n, m\}$ because that is the number of columns in $\mathbf{A}$. This is only important in that we are guranteed enough diagonal positions to hold the r singular values.

b:

Show that $\mathbb{R}^m$ has an orthonormal basis, $\mathbb{R}^n$ has an orthonormal basis, and there exists $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_r \geq 0$ such that following hold:

$$\mathbf{A}\mathbf{v}_i = \begin{cases} \sigma_i \mathbf{u}_i & i = 1, ..., r \\ 0 & i = r+1, ..., m \end{cases}$$
$$\mathbf{A}^T\mathbf{u}_i = \begin{cases} \sigma_i \mathbf{v}_i & i = 1, ..., r \\ 0 & i = r+1, ..., n \end{cases}$$

Answer:
First we will show the the existence of the orthonormal basis and then we will prove the piecewise equations.

Proof.
We know that the matrices $\mathbf{U}$ and $\mathbf{V}^T$ are invertible because they are orthogonal which implies $\mathbf{U}^T = \mathbf{U}^{-1}$, $\mathbf{V}^T = \mathbf{V}^{-1}$. This is then equivalent to saying the columns are linearly independent in both matrices. Because $\mathbf{U}$ and $\mathbf{V}^T$ are of size n and m respectively, their columns then form a basis for $\mathcal{R}^n$ and $\mathcal{R}^m$ respectively. Furthermore, the columns of $\mathbf{U}$ and $\mathbf{V}^T$ are already normalized in the SVD decomposition – thus we have the following:

$$\{\mathbf{u}_1, ..., \mathbf{u}_n\} \text{ basis for } \mathcal{R}^n, \ \{\mathbf{v}_1, ..., \mathbf{v}_m\} \text{ basis for } \mathcal{R}^m$$

Where each vector is orthonormal in its respective set. □

Next we will show that our piecewise equation holds.

Proof.
We will begin with the first piecewise function. Rearranging our singular decomposition of $\mathbf{A}$ using the fact that $\mathbf{V}$ is orthogonal and multiplying on the right we get the following.

$$\mathbf{AV} = \mathbf{U\Sigma}$$

We note that the left and right hand sides are matrices of size $n \times m$, and we can write out the equation for the ith column of each.

$$\mathbf{A}\mathbf{v}_i = \mathbf{U}\mathbf{s}_i$$

Where we note that we have used $\mathbf{v}_i$ to denote the ith column of $\mathbf{V}$, and $\mathbf{s}_i$ to denote the ith column of $\mathbf{\Sigma}$ to differentiate it from the singular values. Because $\mathbf{\Sigma}$ is in a sense diagonal this tells us that only the ith entry of $\mathbf{s}_i$ could be non-zero (assuming $i \leq r$).

$$\implies \mathbf{A}\mathbf{v}_i = \mathbf{u}_i\sigma_i, \ i = 1, ..., r$$

Which follows because we are essentially picking out the ith column of $\mathbf{U}$ and multiplying it by the ith singular value. However, as noted the above equation only holds for the first r columns. After that, each column is the zero vector and thus the result is the zero vector. This gives us our first piecewise equation:

$$\mathbf{A}\mathbf{v}_i = \begin{cases} \sigma_i \mathbf{u}_i & i = 1, ..., r \\ 0 & i = r+1, ..., m \end{cases}$$

Moving on to the second piecewise equation we can again rearrange the decomposition by multiplying on the left by $\mathbf{U^T}$, using its orthogonality property, and taking a transpose.

$$\mathbf{U}^T\mathbf{A} = \mathbf{\Sigma}\mathbf{V}^T \implies \mathbf{A}^T\mathbf{U} = \mathbf{V}\mathbf{\Sigma}^T$$

Like before we will write out the equation for one column of the $m \times n$ matrix on both sides of the equation.

$$\mathbf{A}^T\mathbf{u}_i = \mathbf{V}\mathbf{s}_i$$

This time $\mathbf{s}_i$ is a column of $\mathbf{\Sigma}^T$. If $\mathbf{\Sigma}$ were square then it would be symmetric and transpose would equal the original. That is not necessarily the case here, but the transpose is still "diagonal" in the sense we have defined above. The difference here is that $\mathbf{s}_i$ is of size n, but we still are just picking out the ith column of $\mathbf{V}$ assuming that $i \leq r$. This gives us the following similar result:

$$\implies \mathbf{A}^T\mathbf{u}_i = \mathbf{v}_i\sigma_i, \ i = 1, ..., r$$

For any columns after the rth, the column will be the zero vector and the result will be the zero vector. Thus we have arrived at our second equation:

$$\mathbf{A}^T\mathbf{u}_i = \begin{cases} \sigma_i \mathbf{v}_i & i = 1, ..., r \\ 0 & i = r+1, ..., n \end{cases}$$

□

c:

Argue the following:

$$
\begin{aligned}
&Range(\mathbf{A}) = span\{\mathbf{u}_1, ..., \mathbf{u}_r\}\\
&Null(\mathbf{A}) = span\{\mathbf{v}_{r+1}, ..., \mathbf{v}_m\}\\
&Range(\mathbf{A}^T) = span\{\mathbf{v}_1, ..., \mathbf{v}_r\}\\
&Null(\mathbf{A}^T) = span\{\mathbf{u}_{r+1}, ..., \mathbf{u}_n\}
\end{aligned}
$$

Answer:
We will first define the three terms used above for our matrix. The *Range* of $\mathbf{A}$ is all of the vectors which are a linear combination of its columns – $\{\mathbf{v} : \mathbf{A}\mathbf{x} = \mathbf{v},\ \mathbf{x} \in \mathcal{R}^m\}$. The nullspace of our matrix ($Null(\mathbf{A})$) is given as $\{\mathbf{x} : \mathbf{A}\mathbf{x} = 0, \mathbf{x} \in \mathcal{R}^m\}$. Finally, the span of a set of vectors are all the vectors that are a linear combination of the vectors in the set. With this in mind we will now prove the four relations given above.

Proof.
We begin by showing $Range(\mathbf{A}) = span\{\mathbf{u}_1, ..., \mathbf{u}_r\}$. Let $\mathbf{v}_1$ be in the range of $\mathbf{A}$, or in other words $\mathbf{A}\mathbf{x} = \mathbf{v}_1$ for some $\mathbf{x}$.

$$\mathbf{A}\mathbf{x} = \mathbf{v}_1 \implies \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mathbf{x} = \mathbf{v}_1$$

Let $\mathbf{\Sigma}\mathbf{V}^T\mathbf{x} = \mathbf{z}$ whose last $n - r$ elements, we may note, are zero, as the last $n - r$ rows of $\mathbf{\Sigma}$ are zero vectors.

$$\implies \mathbf{U}\mathbf{z} = \mathbf{v}_1$$

Because the ending elements of $\mathbf{z}$ are zero this means that $\mathbf{v}_1$ is a linear combination of the first r columns of $\mathbf{U}$.

$$\therefore Range(\mathbf{A}) = span\{\mathbf{u}_1, ..., \mathbf{u}_r\}$$

Next, we will show $Range(\mathbf{A}^T) = span\{\mathbf{v}_1, ..., \mathbf{v}_r\}$. Similar to before, let $\mathbf{v}_1$ be in the range of $\mathbf{A}^T$. This means for some $\mathbf{x}$...

$$\mathbf{A}^T\mathbf{x} = \mathbf{v}_1 \implies (\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T)^T\mathbf{x} = \mathbf{v}_1$$

We can distribute the transpose and let $\mathbf{\Sigma}^T\mathbf{U}^T\mathbf{x} = \mathbf{z}$. We will also note that the last $m-r$ elements of $\mathbf{z}$ are zero as the last $m - r$ rows of $\mathbf{\Sigma}^T$ are zero vectors.

$$(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T)^T\mathbf{x} = \mathbf{V}\mathbf{z} = \mathbf{v}_1$$

This means that $\mathbf{v}_1$ is a linear combination of the first r columns of $\mathbf{V}$ which follows because of the structure of $\mathbf{z}$ described above.

$$\therefore Range(\mathbf{A}^T) = span\{\mathbf{v}_1, ..., \mathbf{v}_r\}$$

Having shown the previous two statements it is now much easier to prove the last two. The Fundamental Theorem of Linear Algebra tells us that $Range(\mathbf{A}) \perp Null(\mathbf{A}^T)$ and $Range(\mathbf{A}^{\mathbf{T}}) \perp Null(\mathbf{A})$. We already know that the matrices $\mathbf{U}$ and $\mathbf{V}^{\mathbf{T}}$ are orthogonal matrices. Thus we can use this to build the other spaces based on which vectors are already included.

$$\therefore Null(\mathbf{A}^T) = span\{\mathbf{u}_{r+1}, ..., \mathbf{u}_n\}$$

Which follows from the $Range(\mathbf{A})$.

$$\therefore Null(\mathbf{A}) = span\{\mathbf{v}_{r+1}, ..., \mathbf{v}_m\}$$

Which follows from the $Range(\mathbf{A}^T)$. □

d:

Show that the $Range(\mathbf{A}^T)$ is orthogonal to $Null(\mathbf{A})$.

Answer:

Proof.
Let $x \in Range(\mathbf{A}^T)$ and $y \in Null(\mathbf{A})$ be any two vectors in their respective spaces. Let us then take the inner product of the two...

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{A}^T \mathbf{z}, \mathbf{y} \rangle$$

This follows by definition of the *Range* where $\mathbf{z}$ is some arbitrary non-zero vector. We can then use properties of the inner product to get the following:

$$\langle \mathbf{z}, \mathbf{A}\mathbf{y} \rangle = \langle \mathbf{z}, 0 \rangle = 0$$

This follows because $\mathbf{y}$ is in the Nullspace of $\mathbf{A}$, and the inner product with the zero vector is always zero.

$$\therefore Range(\mathbf{A}^T) \text{ is orthogonal to } Null(\mathbf{A})$$

$\square$