Matrix Computations **Fall 2020**

Lecture 1: Basic Concepts

Lecturer: Prof.Yue Qiu & Prof Ziping Zhao *Scribe: Chenguang Zhang&Zhicheng Wang.*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

1 Vector Space

Definition 1. *A subset $\mathcal{S}$ of $\mathbb{R}^m$ $(\mathbb{C}^m)$ is said to be a* ***subspace*** *if and only if*

1. $x,\ y \in \mathcal{S} \Longrightarrow x + y \in \mathcal{S}$ *closure property for vector addition*
2. $x \in \mathcal{S} \Longrightarrow \alpha x \in \mathcal{S}$ *for* $\forall\ \alpha \in \mathbb{R}^n$ $(\mathbb{C}^n)$ *closure property for scalar multiplication*

Definition 2. *The* ***inner product*** *of two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ is defined as*

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} y_i x_i = \boldsymbol{y}^T \boldsymbol{x}.$$

- $\boldsymbol{x}, \boldsymbol{y}$ *are said to be* ***orthogonal*** *to each other if* $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$*;*
- $\boldsymbol{x}, \boldsymbol{y}$ *are said to be* ***parallel*** *if* $\boldsymbol{x} = \alpha \boldsymbol{y}$ *for some* α*.*

Definition 3. *For a set of vectors $\mathcal{S} = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_r\}$, the subspace*

$$span(\mathcal{S}) = \{\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_r \mathbf{v}_r\},$$

generated by forming all linear combinations of vectors from $\mathcal{S}$ is called the ***space spanned*** *by $\mathcal{S}$.*

Definition 4. *If $\mathcal{X}$ and $\mathcal{Y}$ are subspaces of a vector space V, define the sum of two subspaces by*

$$\mathcal{X} + \mathcal{Y} = \{\mathbf{x} + \mathbf{y} | \mathbf{x} \in \mathcal{X} \text{ and } \mathbf{y} \in \mathcal{Y}\},$$

and if $\mathcal{X} \cap \mathcal{Y} = \mathbf{0}$, $\mathcal{X} + \mathcal{Y} = \mathcal{Z}$, we define the ***direct sum*** *$\mathcal{Z} = \mathcal{X} \oplus \mathcal{Y}$.*

Exercise:

1. For a vector space $\mathcal{V}$, and for $\mathcal{X},\ \mathcal{Y} \subseteq \mathcal{V}$, explain why $span(\mathcal{X} \cup \mathcal{Y}) = span(\mathcal{X}) + span(\mathcal{Y})$.
2. Let $\mathcal{X}$ and $\mathcal{Y}$ be two subspace of a vector space $\mathcal{V}$.
   (a) Prove that the intersection $\mathcal{X} \cap \mathcal{Y}$ is also a subspace of $\mathcal{V}$.
   (b) Show that union $\mathcal{X} \cup \mathcal{Y}$ is only a subspace if $\mathcal{X} \subseteq \mathcal{Y}$ or $\mathcal{Y} \subseteq \mathcal{X}$.
3. For a set of vectors $\mathcal{S} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, prove that span$(\mathcal{S})$ is the intersection of all subspaces that contain $\mathcal{S}$, i.e., prove that span$(\mathcal{S}) = \mathcal{M}$ where $\mathcal{M} := \cap_{s \subseteq \mathcal{V}} \mathcal{V}$ is the intersection of all subspaces that contain $\mathcal{S}$ and $\mathcal{V}$ denotes the subspace containing $\mathcal{S}$.

Solutions:

1. For $\forall\ \boldsymbol{u} \in span(\mathcal{X}) + span(\mathcal{Y}) \iff \exists\ \alpha, \beta,\ s.t. \boldsymbol{u} = \alpha_1 \boldsymbol{x}_1 + \ldots + \alpha_m \boldsymbol{x}_m + \beta_1 \boldsymbol{y}_1 + \ldots + \beta_n \boldsymbol{y}_n = \gamma_1 \boldsymbol{x}_1 + \ldots + \gamma_m \boldsymbol{x}_m + \gamma_{m+1} \boldsymbol{y}_1 + \ldots + \gamma_{m+n} \boldsymbol{y}_n \iff \boldsymbol{u} \in span(\mathcal{X} \cup \mathcal{Y})$.

2. (a) $\forall\ \boldsymbol{u},\ \boldsymbol{v} \in \mathcal{X} \cap \mathcal{Y}$, since $\mathcal{X}$ and $\mathcal{Y}$ are subspace, $\alpha \boldsymbol{u} \in \mathcal{X}$ and $\alpha \boldsymbol{u} \in \mathcal{Y}$, $\alpha \boldsymbol{u} \in \mathcal{X} \cap \mathcal{Y}$, similarly we have $\boldsymbol{u} + \boldsymbol{v} \in \mathcal{X} \cap \mathcal{Y}$. Therefore, $\mathcal{X} \cap \mathcal{Y}$ is a subspace.

   (b) Suppose $\mathcal{X} \not\subseteq \mathcal{Y}$ and $\mathcal{Y} \not\subseteq \mathcal{X}$, $\exists \boldsymbol{u} \in \mathcal{X} - \mathcal{Y},\ \boldsymbol{v} \in \mathcal{Y} - \mathcal{X}$. $\boldsymbol{u} + \boldsymbol{v} \notin \mathcal{X}$, $\boldsymbol{u} + \boldsymbol{v} \notin \mathcal{Y}$ and $\boldsymbol{u} + \boldsymbol{v} \in \mathcal{X} \cup \mathcal{Y}$, a contradiction.

3. The proof consists of two parts:

   - First we prove that $\text{span}(\mathcal{S}) \subset \mathcal{M}$. For any $\mathbf{x} \in \text{span}(\mathcal{S})$, then $\mathbf{x}$ can be linearly represented by $\mathcal{S}$, i.e., $\mathbf{x} = \sum_i \alpha_i \mathbf{v}_i$. For any subspace $\mathcal{V}$ containing $\mathcal{S}$, we must have $\mathbf{x} \in \mathcal{V}$ since subspace $\mathcal{V}$ is closed under addition. Therefore, we have $\mathbf{x} \in \cap_{s \subseteq \mathcal{V}} \mathcal{V} = \mathcal{M}$. To sum up, $\forall \mathbf{x} \in \text{span}(\mathcal{S}) \Rightarrow \mathbf{x} \in \mathcal{M} \Rightarrow \text{span}(\mathcal{S}) \subset \mathcal{M}$.
   - Then we try to prove that $\mathcal{M} \subset \text{span}(\mathcal{S})$. By definition, $\mathcal{M}$ is contained in every subspace which contains $\mathcal{S}$. (The intersection of subspaces is also a subspace.) And since $\text{span}(\mathcal{S})$ is also a subspace, then we have $\mathcal{M} \subset \text{span}(\mathcal{S})$.

   Therefore $\text{span}(\mathcal{S}) = \mathcal{M}$.

2 Linear Independence

Definition 5. *A collection of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^m$ is said to be **linearly independent** if $\sum_{i=1}^n \alpha_i \mathbf{a}_i \neq \mathbf{0}$, for all $\alpha_i \in \mathbb{R}$ with $\alpha_i \neq 0$; and linearly dependent otherwise.*

1. *If $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subset \mathcal{R}^m$ is linearly independent, then any $\mathbf{a}_j$ cannot be a linear combination of the other $\mathbf{a}_i$'s; i.e., $\mathbf{a}_j \neq \sum_{i \neq j} \alpha_i \mathbf{a}_i$ for any α_i's.*

2. *If $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subset \mathcal{R}^m$ is linearly dependent, then there exists an $\mathbf{a}_j$ such that $\mathbf{a}_j$ is a linear combination of the other $\mathbf{a}_i$'s; i.e., $\mathbf{a}_j = \sum_{i \neq j} \alpha_i \mathbf{a}_i$ for some α_i's.*

3. *If $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subset \mathcal{R}^m$ is linearly independent, then $m \geq n$ must hold.*

4. *Let $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subset \mathcal{R}^m$ be a linearly independent vector set. Suppose $\boldsymbol{y} \in span\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$, then the coefficient α for the representation $\boldsymbol{y} = \sum_{i=1}^n \alpha_i \mathbf{a}_i$ is unique; i.e., there does not exist a $\beta \in \mathcal{R}^n$, $\beta \neq \alpha$, such that $\boldsymbol{y} = \sum_{i=1}^n \beta_i \mathbf{a}_i$.*

Definition 6. *A vector subset $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_k}\}$ is called a **maximal linear independent subset** of $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ if:*

1. *$\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_k}\}$ is **linearly independent**;*

2. *$\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_k}\}$ is **not contained** by any other linearly independent subset of $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$.*

Definition 7. *Let $\mathcal{S} \subseteq \mathbb{R}^m$ be a subspace with $\mathcal{S} \neq \{\mathbf{0}\}$. A vector set $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subset \mathcal{R}^m$ is called a **basis** for $\mathcal{S}$ if $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subset \mathcal{R}^m$ is a **linearly independent** and $\mathcal{S} = span\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$.*

Definition 8. *The **dimension** of a nontrivial subspace $\mathcal{S}$ is the **number of elements of a basis for** $\mathcal{S}$. The dimension of of the trivial subspace $\{\mathbf{0}\}$ is defined as 0.*

Theorem 1. *For a nonempty set of vectors $\mathcal{S} = \{\boldsymbol{u}_1, \ldots \boldsymbol{u}_n\}$ in a space $\mathcal{V}$, the following statements are true.*

- *If $\mathcal{S}$ contains a linearly dependent subset, then $\mathcal{S}$ itself must be linearly dependent.*
- *If $\mathcal{S}$ is linearly dependent, then every subset of $\mathcal{S}$ is also linearly independent.*
- *If $\mathcal{S}$ is linearly dependent and if $\mathbf{v} \in \mathcal{V}$, then the **extension set** $\mathcal{S}_{ext} = \mathcal{S} \cup \{\mathbf{v}\}$ is linearly independent if and only if $\mathbf{v} \notin span(\mathcal{S})$.*
- *If $\mathcal{S} \subseteq \mathbb{R}^m$ and if $n > m$, the $\mathcal{S}$ must be linearly dependent.*

Theorem 2. *All bases for $\mathcal{S}$ have the same number of elements; i.e., if $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ and $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ are basis for $\mathcal{S}$ then $m = n$.*

Proof. Suppose $n \neq m$,without loss of generality $n > m$, $\boldsymbol{u}_i = a_{i1}\boldsymbol{v}_1 + \ldots + a_{im}\boldsymbol{v}_m$, $\mathbf{0} = b_1\boldsymbol{u}_1 + \ldots + b_n\boldsymbol{u}_n = \sum_{i=1}^{n} b_i\boldsymbol{u}_i = \sum_{j=1}^{m}\sum_{i=1}^{n} b_i a_{ij}\boldsymbol{v}_j = \sum_{j=1}^{m} c_j\boldsymbol{v}_j$, $A_{i,j} = a_{ij}$, then $A[b_1, \ldots, b_n]^T = [c_1, \ldots, c_m]^T$. Since $n > m$, $\exists\ \boldsymbol{b} \neq \mathbf{0}$, s.t. $A\boldsymbol{b} = \mathbf{0}$ and $b_1\boldsymbol{u}_1 + \ldots + b_n\boldsymbol{u}_n = \mathbf{0}$, a contradiction. □

Theorem 3. *For vector spaces $\mathcal{M}$ and $\mathcal{N}$ such that $\mathcal{M} \subseteq \mathcal{N}$, the following statements are true.*

1. $\dim \mathcal{M} \leq \dim \mathcal{N}$.
2. *If* $\dim \mathcal{M} = \dim \mathcal{N}$, *then* $\mathcal{M} = \mathcal{N}$

Proof. Let $\dim \mathcal{M} = m$ and $\dim \mathcal{N} = n$, and use an indirect argument to prove 1. If it were the case that $m > n$, then there would exist a linearly independent subset of $\mathcal{N}$ (namely, a basis for $\mathcal{M}$) containing more than n vectors. But this is impossible because $\dim \mathcal{N}$ is the size of a maximal independent subset of $\mathcal{N}$. Thus $m \leq n$. Now prove (1). If $m = n$ but $\mathcal{M} \neq \mathcal{N}$, then there exists a vector $\boldsymbol{x}$ such that $\boldsymbol{x} \in \mathcal{N}$ but $\boldsymbol{x} \notin \mathcal{M}$. If $\mathcal{B}$ is a basis for $\mathcal{M}$, then $\boldsymbol{x} \notin span(\mathcal{B})$, and the extension set $\mathcal{E} = \mathcal{B} \cup \{\boldsymbol{x}\}$ s a linearly independent subset of $\mathcal{N}$. But $\mathcal{E}$ contains $m+1 = n+1$ vectors, which is impossible because $\dim \mathcal{N} = n$ is the size of a maximal independent subset of $\mathcal{N}$. Hence $\mathcal{M} = \mathcal{N}$. □

Theorem 4. *If $\mathcal{X}$ and $\mathcal{Y}$ are subspaces of a vector space $\mathcal{V}$, then*

$$\dim(\mathcal{X} + \mathcal{Y}) = \dim \mathcal{X} + \dim \mathcal{Y} - \dim(\mathcal{X} \cap \mathcal{Y}).$$

Proof. The strategy is to construct a basis for $\mathcal{X} + \mathcal{Y}$ and ount the number of vectors it contains. Let $\mathcal{S} = \boldsymbol{z}_1, \ldots, \mathbf{z}_t$ be a basis for $\mathcal{X} \cap \mathcal{Y}$. Since $\mathcal{S} \subseteq \mathcal{X}$ and $\mathcal{S} \subseteq \mathcal{Y}$, there must exist extension vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ and $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ such that

$$\mathcal{B}_x = \boldsymbol{z}_1, \ldots, \mathbf{z}_t, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m = \text{a basis for} \mathcal{X},$$

and

$$\mathcal{B}_x = \boldsymbol{z}_1, \ldots, \mathbf{z}_t, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n = \text{a basis for} \mathcal{Y}.$$

We know from exercise 1 that $\mathcal{B} = \mathcal{B}_x \cup \mathcal{B}_y$ spans $\mathcal{X} + \mathcal{Y}$, and we wish show that $\mathbb{B}$ is linearly independent. If

$$\sum_{i=1}^{t} \alpha_i \mathbf{z}_i + \sum_{j=1}^{m} \beta_j \boldsymbol{x}_j + \sum_{k=1}^{n} \gamma_k \boldsymbol{y}_k = \mathbf{0},$$

then

$$\sum_{k=1}^{n} \gamma_k \boldsymbol{y}_k = -(\sum_{i=1}^{t} \alpha_i \mathbf{z}_i + \sum_{j=1}^{m} \beta_j \boldsymbol{x}_j) \in \mathcal{X}.$$

Since it is also true that $\sum_{k=1}^{n} \gamma_k \boldsymbol{y}_k \in \mathcal{Y}$, we have that $\sum_{k=1}^{n} \gamma_k \boldsymbol{y}_k \in \mathcal{X} \cap \mathcal{Y}$, and hence there must exist scalars δ_i such that

$$\sum_{k=1}^{n} \gamma_k \boldsymbol{y}_k = \sum_{i=1}^{t} \delta_i \mathbf{z}_i \quad \text{or, equivalently,} \quad \sum_{k=1}^{n} \gamma_k \boldsymbol{y}_k - \sum_{i=1}^{t} \delta_i \mathbf{z}_i = 0.$$

Since $\mathcal{B}$ is an independent set, it follows that all of the γ_k's (as well as all δ_i's) are zero. But $\mathcal{B}_x$ s also an independent set, so the only way this can hold is for all of the α_i's as well as all of the β_j's to be zero. Therefore, the only possible solution for the α's, β's and γ's in the homogeneous equation is the trivial solution, and thus $\mathcal{B}$ is linearly independent. Since $\mathcal{B}$ is an independent spanning set, it is a basis for $\mathcal{X} + \mathcal{Y}$ and, consequently,

$$\dim(\mathcal{X} + \mathcal{Y}) = t + m + n = (t + m) + (t + n) - t = \dim \mathcal{X} + \dim \mathcal{Y} + \dim(\mathcal{X} \cap \mathcal{Y}).$$

□

Theorem 5. *Every subspace of $\mathcal{V}$ is part of a direct sum equal to $\mathcal{V}$:*

Suppose $\mathcal{V}$ is finite-dimensional and $\mathcal{U}$ is a subspace of $\mathcal{V}$. Then there is a subspace $\mathcal{W}$ of $\mathcal{V}$ such that $\mathcal{V} = \mathcal{V} \oplus \mathcal{W}$.

Proof. Proof Because $\mathcal{V}$ is finite-dimensional, so is $\mathcal{U}$. Thus there is a basis $u_1, ..., u_m$ of $\mathcal{U}$. Of course $u_1, ..., u_m$ is a linearly independent list of vectors in $\mathcal{V}$. Hence this list can be extended to a basis $u_1, ..., u_m, w_1, ..., w_n$ of $\mathcal{V}$. Let $\mathcal{W} = span(w_1, ..., w_n)$. To prove that $\mathcal{V} = \mathcal{V} \oplus \mathcal{W}$ by Def. 4 we need only show that

$$\mathcal{U} \cap \mathcal{W} = \mathbf{0} \quad \textbf{and} \quad \mathcal{U} + \mathcal{W} = \mathcal{V}$$

To prove the first equation above, suppose $v \in \mathcal{V}$. Then, because the list $u_1, ..., u_m, w_1, ..., w_n$ spans $\mathcal{V}$, there exist $a_1, ..., a_m, b_1, ..., b_n \in R$ such that

$$v = a_1 u_1 + ... + a_m u_m + b_1 w_1 + ... + b_m w_n.$$

In other words, we have $v = u + w$ where $u \in \mathcal{U}$ and $w \in \mathcal{W}$ are defined as above. Thus $v \in \mathcal{U} \oplus \mathcal{W}$, completing the proof that $\mathcal{V} = \mathcal{U} \cap \mathcal{W}$. To show that $\mathcal{U} + \mathcal{W} = \mathcal{V}$, suppose $v \in \mathcal{U} \oplus \mathcal{W}$. Then there exist scalars $a_1, ..., a_m, b_1, ..., b_n \in R$ such that

$$v = a_1 u_1 + ... + a_m u_m = b_1 w_1 + ... + b_m w_n.$$

Thus

$$a_1 u_1 + ... + a_m u_m - b_1 w_1 - ... - b_m w_n = 0$$

Because $u_1, ..., u_m, w_1, ..., w_n$ is linearly independent, this implies that $a_1 = ... = a_m = b_1 = ... = b_n = 0$. Thus $v = 0$, completing the proof that $\mathcal{U} \cap \mathcal{W} = \mathbf{0}$

□

Fundamental Subspaces—Dimension and Bases

For an $m \times n$ matrix of real numbers such that $rank(\mathbf{A}) = r$,

- $\dim R(\mathbf{A}) = r$,
- $\dim N(\mathbf{A}) = n - r$,
- $\dim R(\mathbf{A}^T) = r$,
- $\dim N(\mathbf{A}^T) = m - r$,

Exercise:

1. If $\mathcal{S}_r = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is a linearly independent subset of an n -dimensional space $\mathcal{V}$, where $r < n$, explain why it must be possible to find extension vectors $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ from $\mathcal{V}$ such that

$$\mathcal{S}_n = \{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$$

is a basis for $\mathcal{V}$.

Solutions:

1. $r < n$ means that *span*$(\mathcal{S}_r) \neq \mathcal{V}$, and hence there exists a vector $\mathbf{v}_{r+1} \in \mathcal{V}$ such that $\mathbf{v}_{r+1} \notin \mathcal{S}_r$. The extension set $\mathcal{S}_{r+1} = \mathcal{S}_r \cup \{\mathbf{v}_{r+1}\}$ is an independent subset of $\mathcal{V}$ containing $r+1$ vectors. Repeating this process generates independent subsets $\mathcal{S}_{r+2}, \mathcal{S}_{r+3}, \ldots,$ and eventually leads to a maximal independent subset $\mathcal{S}_n \subset \mathcal{V}$ containing n vectors.

3 Linear System

Definition 9. *The **range of a matrix** $\mathbf{A} \in \mathcal{R}^{m\times n}$ is defined to be the subspace $R(\mathbf{A})$ of $\mathcal{R}^m$ that is generated by the range of $f(\boldsymbol{x}) = \mathbf{A}\boldsymbol{x}$. That is,*

$$R(\mathbf{A}) = \{\mathbf{A}\boldsymbol{x} | \boldsymbol{x} \in \mathcal{R}^n\} \subseteq \mathcal{R}^m.$$

Similarly, the range of $\mathbf{A}^T$ is the subspace of $\mathcal{R}^n$ defined by

$$R(\mathbf{A}^T) = \{\mathbf{A}^T\boldsymbol{y} | \boldsymbol{y} \in \mathcal{R}^m\} \subseteq \mathcal{R}^n.$$

*Because $R(\mathbf{A})$ is the set of all "images" of vectors $\boldsymbol{x} \in \mathcal{R}^m$ under transformation by $\mathbf{A}$ some people call $R(\mathbf{A})$ the **image space** of $\mathbf{A}$.*

1. *$R(\mathbf{A})$ is the space spanned by the columns of $\mathbf{A}$ (column space).*
2. *$R(\mathbf{A}^T)$ is the space spanned by the rows of $\mathbf{A}$ (row space).*
3. *$\boldsymbol{b} \in R(\mathbf{A}) \Longleftrightarrow \boldsymbol{b} = \mathbf{A}\boldsymbol{x}$ for some $\boldsymbol{x}$.*
4. *$\boldsymbol{a} \in R(\mathbf{A}^T) \Longleftrightarrow \boldsymbol{a}^T = \boldsymbol{y}^T\mathbf{A}$ for some $\boldsymbol{y}^T$.*

Definition 10. *The **null space (nullspace)** or **kernel space** of $\mathbf{A} \in \mathcal{R}^{m\times n}$ is defined as*

$$N(\mathbf{A}) = \{\boldsymbol{x} \in \mathcal{R}^n | \mathbf{A}\boldsymbol{x} = \mathbf{0}\}.$$

Definition 11. *The **rank** of a matrix $\mathbf{A} \in \mathcal{R}^{m\times n}$, denoted by rank$(\mathbf{A})$, is defined as the number of of elements of a maximal linearly independent subset of $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$.*

Definition 12. *$\mathbf{A} \in \mathcal{R}^{m\times n}$ is said to have*

- ***full column rank** if the columns of $\mathbf{A}$ are linearly independent (more precisely, the collection of all columns of $\mathbf{A}$ is linearly independent)*

$$\mathbf{A} \text{ being of full column rank} \Longleftrightarrow m \geq n,\ \text{rank}(\mathbf{A}) = n;$$

- ***full row rank** full row rank $\mathbf{A}$ are linearly independent*

$$\mathbf{A} \text{ being of full row rank} \Longleftrightarrow m \leq n,\ \text{rank}(\mathbf{A}) = m;$$

- ***full rank*** *if* $rank(\mathbf{A}) = \min\{m, n\}$*; i.e., it has either full column rank or full row rank;*
- ***rank deficient*** *if* $rank(\mathbf{A}) < \min\{m, n\}$

Theorem 6. $\dim R(\mathbf{A}) + \dim N(\mathbf{A}) = n$*, for all* $m \times n$ *matrices.*

Theorem 7. *If* $\mathbf{A}$ *is an* $m \times n$ *matrix then*

1. $N(\mathbf{A}) = \{\mathbf{0}\}$ *if and only if* $rank(\mathbf{A}) = n$*;*
2. $N(\mathbf{A}^T) = \{\mathbf{0}\}$ *if and only if* $rank(\mathbf{A}) = m$*;*

Exercise:

1. Determine spanning sets for $R(\mathbf{A})$ and $N(\mathbf{A})$, where
$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 & 3 \\ 2 & 4 & 1 & 3 \\ 3 & 6 & 1 & 4 \end{bmatrix}.$$
2. If $\mathbf{A}$ is $m \times n$ and $\mathbf{B}$ is $n \times p$ then,
   (a) $rank(\boldsymbol{AB}) \leq \min\{rank(\mathbf{A}), rank(\mathbf{B})\}$,
   (b) $rank(\mathbf{A}) + rank(\mathbf{B}) - n \leq rank(\boldsymbol{AB})$.

Solution:

1. Reducing $\mathbf{A}$ to any row echelon form $\mathbf{U}$ provides the solution—the basic columns in $\mathbf{A}$ correspond to the pivotal positions in $\mathbf{U}$, and the nonzero rows of $\mathbf{U}$ span the row space of $\mathbf{A}$.
$$\begin{bmatrix} 1 & 2 & 2 & 3 \\ 2 & 4 & 1 & 3 \\ 3 & 6 & 1 & 4 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 2 & 3 \\ 0 & 0 & -3 & -3 \\ 0 & 0 & -5 & -5 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
Therefore, $R(\mathbf{A}) = span\{\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}\}$ and $N(\mathbf{A}) = span\{\begin{pmatrix} 1 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ -2 \end{pmatrix}\}$
2. (a) Let $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p], \mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p]$.
   - Let $\mathbf{AB} = [A\mathbf{b}_1, \mathbf{Ab}_2, \ldots, \mathbf{Ab}_p]$. $\forall j \in [p], A\mathbf{b}_j$ is a linear combination of $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$, so $\mathbf{Ab}_j \in \mathsf{span}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$, i.e., $\{\mathbf{Ab}_1, \mathbf{Ab}_2, \ldots, \mathbf{Ab}_p\} \subset \mathsf{span}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$, $\mathsf{span}(\mathbf{Ab}_1, \mathbf{Ab}_2, \ldots, \mathbf{Ab}_p) \subset \mathsf{span}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$, we then get $\mathsf{rank}(\mathbf{AB}) \leq \mathsf{rank}(\mathbf{A})$.
   - Since $\mathsf{rank}(\mathbf{AB}) = \mathsf{rank}(\mathbf{B}^T\mathbf{A}^T)$, applying the same strategy, we can get $\mathsf{rank}(\mathbf{B}^T\mathbf{A}^T) \leq \mathsf{rank}(\mathbf{B}^T) = \mathsf{rank}(\mathbf{B})$.

   Therefore, $\mathsf{rank}(\mathbf{AB}) \leq \min\{\mathsf{rank}(\mathbf{A}), \mathsf{rank}(\mathbf{B})\}$.

   (b) $\begin{bmatrix} \mathbf{A} & \mathbf{AB} \\ \boldsymbol{E}_n & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{E}_n & \mathbf{B} \\ \mathbf{0} & \boldsymbol{E}_p \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \boldsymbol{E}_n & \mathbf{B} \end{bmatrix}$, since $\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \boldsymbol{E}_n & \mathbf{B} \end{bmatrix}$ is invertible, then $rank(\mathbf{A}) + rank(\mathbf{B}) \leq rank(\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \boldsymbol{E}_n & \mathbf{B} \end{bmatrix}) = rank(\begin{bmatrix} \mathbf{A} & \mathbf{AB} \\ \boldsymbol{E}_n & \mathbf{0} \end{bmatrix}) = rank(\begin{bmatrix} \mathbf{0} & \boldsymbol{AB} \\ \boldsymbol{E}_n & \mathbf{0} \end{bmatrix}) = rank(\mathbf{AB}) + rank(\boldsymbol{E}_n) = rank(\mathbf{AB}) + n$. Therefore, $rank(\mathbf{AB}) - n \leq rank(\mathbf{AB})$.

4 Orthogonal Complements and Projections

Theorem 8. *Direct sum of a subspace and its orthogonal complement:*
Suppose U is a finite-dimensional subspace of V. Then

$$V = U \oplus U^{\perp}$$

Definition 13. *Suppose U is a finite-dimensional subspace of V. The orthogonal projection of V onto U is the operator $P_U \in \mathcal{L}(V)$ defined as follows:*
For $v \in V$, write $v = u + w$, where $u \in U$ and $w \in U$. Then $P_U(v) = u$.

Theorem 9. *Minimizing the distance to a subspace: Suppose U is a finite-dimensional subspace of V, $v \in V$, and $u \in U$. Then*

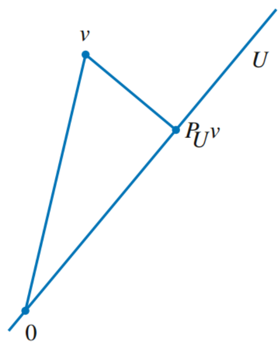
$$||v - P_U(v)|| \leq ||v - u||.$$

Furthermore, the inequality above is an equality if and only if $u = P_U(v)$.

Proof.

$$\begin{aligned} ||v - P_U(v)||^2 &\leq ||v - P_U(v)||^2 + ||P_U(v) - u||^2 \\ &= ||(v - P_U(v)) + (P_U(v) - u)||^2 \\ &= ||v - u||^2, \end{aligned}$$

where the first line above holds because $||v - P_U(v)||^2 \geq 0$, the second line above comes from the Pythagorean Theorem $[v - P_U(v) \perp u, P_U(v)]$, and the third line above holds by simple algebra. Taking square roots gives the desired inequality. Our inequality above is an equality if and only if $||v - P_U(v)|| = 0$, which happens if and only if $v = P_U(v)$. □



$P_U v$ *is the closest point in* U *to* v.

References

[1] Horn, Roger A., and Charles R. Johnson. Matrix analysis. Cambridge university press, 2012.

[2] Meyer, Carl D. Matrix analysis and applied linear algebra. Vol. 71. Siam, 2000.

[3] Axler, Sheldon. Linear algebra done right. springer, 2015.