Fundamental subspaces

COMS 3251 Fall 2022 (Daniel Hsu)

1 Column space

Recall that the *column space* of an $m \times n$ matrix A, written $\mathsf{CS}(A)$, is the span of the columns of A. Moreover, we have seen that $\mathsf{CS}(A)$ is a subspace of $\mathbb{R}^m$, and that $\mathsf{CS}(A) = \mathsf{CS}(C)$, where C comes from the CR factorization of A. Since the columns of C are linearly independent, they form a basis for $\mathsf{CS}(A)$, and hence the dimension of $\mathsf{CS}(A)$ is equal to the number of columns of C, which is $\operatorname{rank}(A)$.

2 Nullspace

2.1 Definition and basic properties

The *nullspace* of an $m \times n$ matrix A, written $\mathsf{NS}(A)$[1] , is the set of vectors $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{0}$, i.e., the solution set for the homogeneous system of linear equations with coefficient matrix A. It is also the set of vectors that are "nullified" by the linear transformation $T_A \colon \mathbb{R}^n \to \mathbb{R}^m$ given by $T_A(\mathbf{v}) = A\mathbf{v}$.

Proposition 1. *The nullspace of an $m \times n$ matrix A is a subspace of $\mathbb{R}^n$.*

Proof. It is clear from the definition that $\mathsf{NS}(A) \subseteq \mathbb{R}^n$. So it suffices to verify that $\mathsf{NS}(A)$ satisfies SS1, SS2, and SS3. First, suppose $\mathbf{u}, \mathbf{v} \in \mathsf{NS}(A)$. Then for any $c \in \mathbb{R}$, linearity of matrix-vector multiplication guarantees

$$A(c\,\mathbf{u} + \mathbf{v}) \;=\; c\,(A\mathbf{u}) + (A\mathbf{v}) \;=\; c\,\mathbf{0} + \mathbf{0} \;=\; \mathbf{0},$$

so $c\,\mathbf{u} + \mathbf{v} \in \mathsf{NS}(A)$ as well. This verifies SS1 and SS2. Second, $\mathbf{0} \in \mathsf{NS}(A)$ since $A\mathbf{0} = \mathbf{0}$. This verifies SS3. □

The following proposition follows immediately from definitions.

Proposition 2. *The columns of an $m \times n$ matrix A are linearly independent if and only if $\mathsf{NS}(A) = \{\mathbf{0}\}$.*

[1] Some texts refer to the nullspace of A by $\mathrm{N}(A)$.

Proposition 2 says that $A\mathbf{x} = \mathbf{0}$ has exactly one solution—namely, $\mathbf{x} = \mathbf{0}$—if and only if the columns of A are linearly independent.

So if the columns of A are not linearly independent, then there is a "non-trivial" (i.e., not just $\{\mathbf{0}\}$) subspace of solutions to $A\mathbf{x} = \mathbf{0}$. To truly "solve" $A\mathbf{x} = \mathbf{0}$, we need to characterize the entire subspace of solutions, and we can do so by determining a basis for the subspace. So our goal will be to find a basis for $\mathsf{NS}(A)$.

2.2 Basis for the nullspace

The CR factorization of A gives one way to determine a basis for $\mathsf{NS}(A)$. Recall that if $d = \operatorname{rank}(A)$, then in the CR factorization $A = CR$, the matrix $C = [\mathbf{c}_1, \dots, \mathbf{c}_d]$ contains a maximal subset of linearly independent columns of A, and $R = [\mathbf{r}_1, \dots, \mathbf{r}_n]$ is a matrix in RREF, with no all-zeros rows, that reveals how to reconstruct every column of $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ using linear combinations of those in C. Let $\mathsf{PV}(R) = \{i_1, \dots, i_d\} \subseteq \{1, \dots, n\}$ be the indices of columns that contain a pivot in R (indices of pivot variables), with $i_1 < \cdots < i_d$, and let $\mathsf{FV}(R) = \{1, \dots, n\} \setminus \mathsf{PV}(R)$ (indices of free variables). Note that if $k \in \mathsf{FV}(R)$ and $\mathbf{r}_k = (r_{1,k}, \dots, r_{d,k})$, then

$$\mathbf{a}_k \;=\; C\mathbf{r}_k \;=\; \sum_{j=1}^{d} r_{j,k}\mathbf{a}_{i_j}.$$

We define a <u>*special solution*</u> $\mathbf{s}_k$ to $A\mathbf{x} = \mathbf{0}$ for each $k \in \mathsf{FV}(R)$ as follows.

1. The kth component of $\mathbf{s}_k$ is equal to 1.
2. For each $j \in \{1, \dots, d\}$, the i_jth component of $\mathbf{s}_k$ is equal to $-r_{j,k}$.
3. All other components of $\mathbf{s}_k$ are equal to 0.

Therefore,

$$A\mathbf{s}_k \;=\; \mathbf{a}_k - \sum_{j=1}^{d} r_{j,k}\mathbf{a}_{i_j} \;=\; \mathbf{0},$$

so $\mathbf{s}_k$ indeed solves $A\mathbf{x} = \mathbf{0}$, i.e., $\mathbf{s}_k \in \mathsf{NS}(A)$.

Example. Consider the matrix

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 4 & 4 \\ 3 & 6 & 5 & 4 \end{bmatrix}.$$

Its CR factorization $A = CR$ is given by

$$C = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_3 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 3 & 5 \end{bmatrix}, \qquad R = \begin{bmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

We have $\mathsf{PV}(R) = \{1, 3\}$ and $\mathsf{FV}(R) = \{2, 4\}$. The special solutions to $A\mathbf{x} = \mathbf{0}$ are

$$\mathbf{s}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{s}_4 = \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}.$$ ∎

Proposition 3. *For any matrix A, the set of special solutions to $A\mathbf{x} = \mathbf{0}$ is a basis for $\mathsf{NS}(A)$.*

Proof. Let $\mathcal{S} = \{\mathbf{s}_k : k \in \mathsf{FV}(R)\}$ denote the set of special solutions to $A\mathbf{x} = \mathbf{0}$, where R is the matrix in RREF obtained from the CR factorization of A. First observe that $\mathcal{S}$ is linearly independent, since for each $\ell \in \mathsf{FV}(R)$, the ℓth component of $\mathbf{s}_k$ is non-zero if and only if $\ell = k$,

So it remains to show that $\operatorname{span}(\mathcal{S}) = \mathsf{NS}(A)$. Consider any $\mathbf{x} = (x_1, \dots, x_n) \in \mathsf{NS}(A)$, and now define $\mathbf{y} = (y_1, \dots, y_n)$ by

$$\mathbf{y} = \mathbf{x} - \sum_{k \in \mathsf{FV}(R)} x_k \, \mathbf{s}_k.$$

Observe that, by linearity of matrix-vector multiplication and the facts that $\mathbf{x} \in \mathsf{NS}(A)$ and $\mathbf{s}_k \in \mathsf{NS}(A)$ for each $k \in \mathsf{FV}(R)$,

$$\begin{aligned} A\mathbf{y} &= A\mathbf{x} - \sum_{k \in \mathsf{FV}(R)} x_k \, (A\mathbf{s}_k) \\ &= A\mathbf{x} - \sum_{k \in \mathsf{FV}(R)} x_k \, \mathbf{0} \\ &= A\mathbf{x} \\ &= \mathbf{0}. \end{aligned}$$

Since the kth component of $\mathbf{s}_k$ is 1 for each $k \in \mathsf{FV}(R)$, it follows that $y_k = 0$ for each $k \in \mathsf{FV}(R)$. So $A\mathbf{y}$ is a linear combination of $\{\mathbf{a}_k : k \in \mathsf{PV}(R)\}$, which is linearly independent. Since we showed $A\mathbf{y} = \mathbf{0}$, it must be that $\mathbf{y} = \mathbf{0}$, so

$$\mathbf{0} \ = \ \mathbf{x} - \sum_{k \in \mathsf{FV}(R)} x_k \mathbf{s}_k .$$

This shows $\mathbf{x} \in \operatorname{span}(\mathcal{S})$; hence $\operatorname{span}(\mathcal{S}) = \mathsf{NS}(A)$. □

The basis for $\mathsf{NS}(A)$ has cardinality $n - d$: one vector per free variable. So $\mathsf{NS}(A)$ has dimension equal to $n - d = n - \operatorname{rank}(A)$. We have thus shown the following theorem.

Theorem 1 (Dimension Theorem)**.** *For any matrix A with n columns,*

$$\operatorname{rank}(A) + \dim(\mathsf{NS}(A)) \ = \ n.$$

The dimension of the nullspace of A is called the *nullity* of A, so Theorem 1 is also called the "Rank-Nullity Theorem".

2.3 Solving general systems of linear equations

Elimination is able to find a solution to an arbitrary system of linear equations $A\mathbf{x} = \mathbf{b}$ (assuming one exists, which we do for the remainder of this section). If the columns of A are linearly independent, then that solution is unique. However, if the columns of A are not linearly independent, then there are infinitely-many solutions. This is because if $\mathbf{x}_{\mathsf{particular}}$ is a solution, then so is $\mathbf{x}_{\mathsf{particular}} + \mathbf{z}$ for every $\mathbf{z} \in \mathsf{NS}(A)$:

$$A(\mathbf{x}_{\mathsf{particular}} + \mathbf{z}) \ = \ A\mathbf{x}_{\mathsf{particular}} + A\mathbf{z} \ = \ \mathbf{b} + \mathbf{0} \ = \ \mathbf{b}.$$

Given any "particular solution" $\mathbf{x}_{\mathsf{particular}}$ to $A\mathbf{x} = \mathbf{b}$, it turns out (as we'll see in Proposition 4) that the entire solution set is

$$\{\mathbf{x}_{\mathsf{particular}}\} + \mathsf{NS}(A).$$

Above, we are using *sumset* notation $\mathcal{S} + \mathcal{T} = \{\mathbf{s} + \mathbf{t} : \mathbf{s} \in \mathcal{S}, \mathbf{t} \in \mathcal{T}\}$. The only thing unsatisfying about this description is that writing "$\mathsf{NS}(A)$" is not very explicit about what vectors are in $\mathsf{NS}(A)$.

But if $\mathbf{z} \in \mathsf{NS}(A)$, then it is a linear combination of the special solutions $\{\mathbf{s}_k : k \in \mathsf{FV}(R)\}$, as per Proposition 3. So, to give an explicit description of the entire solution set for $A\mathbf{x} = \mathbf{0}$, we first find any "particular solution" $\mathbf{x}_{\mathsf{particular}}$ to $A\mathbf{x} = \mathbf{b}$, then we find the set of special solutions $\{\mathbf{s}_k : k \in \mathsf{FV}(R)\}$, and finally we declare that the solution set is

$$\{\mathbf{x}_{\mathsf{particular}}\} + \operatorname{span}(\{\mathbf{s}_k : k \in \mathsf{FV}(R)\}),$$

or equivalently,

$$\left\{ \mathbf{x}_{\mathsf{particular}} + \sum_{k \in \mathsf{FV}(R)} c_k \, \mathbf{s}_k : c_k \in \mathbb{R} \text{ for all } k \in \mathsf{FV}(R) \right\}.$$

Proposition 4. *Assume the system of linear equations $A\mathbf{x} = \mathbf{b}$ has at least one solution, and let $\mathbf{x}_{\mathsf{particular}}$ be any such solution. Then the solution set for $A\mathbf{x} = \mathbf{b}$ is*

$$\{\mathbf{x}_{\mathsf{particular}}\} + \operatorname{span}(\mathcal{S}),$$

where $\mathcal{S} = \{\mathbf{s}_k : k \in \mathsf{FV}(R)\}$ is the set of special solutions to $A\mathbf{x} = \mathbf{0}$.

Proof. Suppose $\mathbf{x}$ satisfies $A\mathbf{x} = \mathbf{b}$. Then

$$A(\mathbf{x} - \mathbf{x}_{\mathsf{particular}}) \;=\; A\mathbf{x} - A\mathbf{x}_{\mathsf{particular}} \;=\; \mathbf{b} - \mathbf{b} = \mathbf{0},$$

which implies $\mathbf{x} - \mathbf{x}_{\mathsf{particular}} \in \mathsf{NS}(A)$. Since $\operatorname{span}(\mathcal{S}) = \mathsf{NS}(A)$ by Proposition 3, it follows that $\mathbf{x} = \mathbf{x}_{\mathsf{particular}} + (\mathbf{x} - \mathbf{x}_{\mathsf{particular}}) \in \{\mathbf{x}_{\mathsf{particular}}\} + \operatorname{span}(\mathcal{S})$ as claimed. □

Example. Consider the system of linear equations $A\mathbf{x} = \mathbf{b}$ where

$$A \;=\; \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 4 & 4 \\ 3 & 6 & 5 & 4 \end{bmatrix}, \quad \mathbf{b} \;=\; \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}.$$

Using Elimination, we transform the augmented matrix to one in which the coefficient matrix is in RREF:

$$\left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 3 \\ 2 & 4 & 4 & 4 & 2 \\ 3 & 6 & 5 & 4 & 1 \end{array}\right] \quad \longrightarrow \quad \left[\begin{array}{cccc|c} 1 & 2 & 0 & -2 & -3 \\ 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

To obtain a solution, we can set the values of the free variables to 0, and assign the values of the pivot variables accordingly:

$$\mathbf{x}_{\text{particular}} = \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \end{bmatrix}.$$

The special solutions to $A\mathbf{x} = \mathbf{0}$ are

$$\mathbf{s}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{s}_4 = \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}.$$

So the solution set for $A\mathbf{x} = \mathbf{b}$ is

$$\left\{ \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + c_4 \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix} : (c_2, c_4) \in \mathbb{R}^2 \right\}.$$

∎

3 Row space

View a $m \times n$ matrix A as a stack of m row vectors. These row vectors are just n-vectors that are lying down horizontally. Suppose these rows are turned back into columns, and we arrange these columns side-by-side in the same order as they were in as rows in A. The result is an $n \times m$ matrix, called the *transpose* of A, written $A^{\intercal}$.

The *row space* of A is the span of the columns of $A^{\intercal}$, written $\mathsf{CS}(A^{\intercal})$, and it is a subspace of $\mathbb{R}^n$.[2] Our analysis of the CR factorization shows that maximum number of linearly independent rows is the same as the maximum number of linearly independent columns. So the dimension of $\mathsf{CS}(A^{\intercal})$ is the same as the dimension of $\mathsf{CS}(A)$, which is $\operatorname{rank}(A)$. A basis for $\mathsf{CS}(A^{\intercal})$ is provided by the rows of R from the CR factorization of A (after turning these rows back into columns).

[2] It might make more sense to define the row space of A to be the set of all linear combinations of the rows of A. But that set is not a subspace of $\mathbb{R}^n$, since n-vectors are column vectors, not row vectors.

Aside: transpose of a composition. A linear combination of the rows of A can be expressed as $\mathbf{y}^\top A$ for some m-vector $\mathbf{y}$. If such a row vector is turned into a column (say, so that it belongs to $\mathsf{CS}(A^\top)$), then we obtain a linear combination of the columns of $A^\top$, written as $A^\top \mathbf{y}$.

In general, for matrices Y and A such that the multiplication $Y^\top A$ is valid, then converting the rows of $Y^\top A$ into columns is obtained by $(Y^\top A)^\top = A^\top Y$.

4 Left nullspace

The *left nullspace* of an $m \times n$ matrix A, written $\mathsf{NS}(A^\top)$, is the nullspace of $A^\top$. We can view $\mathsf{NS}(A^\top)$ as a subspace of $\mathbb{R}^m$. By interchanging the roles of rows and columns, we find that $\dim(\mathsf{NS}(A^\top)) = m - \operatorname{rank}(A)$, i.e.,

$$\operatorname{rank}(A) + \dim(\mathsf{NS}(A^\top)) \;=\; m.$$

5 Fundamental subspaces

We defined four *fundamental subspaces* associated with an $m \times n$ matrix A:

1. column space $\mathsf{CS}(A)$, a subspace of $\mathbb{R}^m$ of dimension $\operatorname{rank}(A)$;
2. row space $\mathsf{CS}(A^\top)$, a subspace of $\mathbb{R}^n$ of dimension $\operatorname{rank}(A)$;
3. nullspace $\mathsf{NS}(A)$, a subspace of $\mathbb{R}^n$ of dimension $n - \operatorname{rank}(A)$;
4. left nullspace $\mathsf{NS}(A^\top)$, a subspace of $\mathbb{R}^m$ of dimension $m - \operatorname{rank}(A)$.

We can define analogous fundamental subspaces for general linear transformations $T \colon \mathbb{V} \to \mathbb{W}$ between general vector spaces $\mathbb{V}$ and $\mathbb{W}$. Here, we just consider two of them:

- The *image* of T (a.k.a. *range*), written $\operatorname{im}(T)$, is defined to be

  $$\operatorname{im}(T) \;=\; \{T(\mathbf{v}) : \mathbf{v} \in \mathbb{V}\},$$

  and it is a subspace of $\mathbb{W}$.

- The *kernel* of T, written $\ker(T)$, is defined to be

  $$\ker(T) \;=\; \{\mathbf{v} \in \mathbb{V} : T(\mathbf{v}) = \mathbf{0}\},$$

  and it is a subspace of $\mathbb{V}$.

We have the following generalization of the Dimension Theorem.

Theorem 2 (Dimension Theorem, General Version)**.** *Let $\mathbb{V}$ be a finite dimensional vector space, and $\mathbb{W}$ be another vector space. If $T\colon \mathbb{V} \to \mathbb{W}$ is linear, then*

$$\dim(\operatorname{im}(T)) + \dim(\ker(T)) \;=\; \dim(\mathbb{V}).$$

A General proof of the Dimension Theorem

Proof of Theorem 2. Let $n = \dim(\mathbb{V})$, and assume it is finite. Let $T\colon \mathbb{V} \to \mathbb{W}$ be linear. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ be a basis for $\ker(T)$, where $k = \dim(\ker(T))$. By the Basis Completion Theorem, there are vectors $\mathbf{v}_{k+1}, \dots \mathbf{v}_n \in \mathbb{V}$ such that $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis for $\mathbb{V}$.

We claim that $\mathcal{S} = \{T(\mathbf{v}_{k+1}), \dots, T(\mathbf{v}_n)\}$ is a basis for $\operatorname{im}(T)$ with $|\mathcal{S}| = n - k$. To do this, we need to show:

1. $\operatorname{span}(\mathcal{S}) = \operatorname{im}(T)$,
2. $T(\mathbf{v}_i) \neq T(\mathbf{v}_j)$ for all $k+1 \leq i < j \leq n$, and
3. $\mathcal{S}$ is linearly independent.

We first show that $\operatorname{span}(\mathcal{S}) = \operatorname{im}(T)$. We know, by definition and linearity, that $\operatorname{im}(T) = \operatorname{span}(\{T(\mathbf{v}_1), \dots, T(\mathbf{v}_n)\})$ since $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis for $\mathbb{V}$. However, $T(\mathbf{v}_1) = \cdots = T(\mathbf{v}_k) = \mathbf{0}$ since $\mathbf{v}_1, \dots, \mathbf{v}_k \in \ker(T)$. So, we have This implies that

$$\begin{aligned} \operatorname{im}(T) &= \operatorname{span}(\{T(\mathbf{v}_1), \dots, T(\mathbf{v}_n)\}) \\ &= \operatorname{span}(\{\mathbf{0}, T(\mathbf{v}_{k+1}), \dots, T(\mathbf{v}_n)\}) \\ &= \operatorname{span}(\mathcal{S}), \end{aligned}$$

where the final step follows by the Removal Theorem.

Next, we show that $T(\mathbf{v}_i) \neq T(\mathbf{v}_j)$ for all $k+1 \leq i < j \leq n$. Suppose for sake of contradiction that $T(\mathbf{v}_i) = T(\mathbf{v}_j)$ for some $k+1 \leq i < j \leq n$. Then we have $T(\mathbf{v}_i) - T(\mathbf{v}_j) = T(\mathbf{v}_i - \mathbf{v}_j) = \mathbf{0}$, i.e., $\mathbf{v}_i - \mathbf{v}_j \in \ker(T)$. Since $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a basis for $\ker(T)$, there are scalars $c_1, \dots, c_k$ such that

$$c_1 \mathbf{v}_1 + \cdots + c_k \mathbf{v}_k = \mathbf{v}_i - \mathbf{v}_j.$$

In other words, the following linear combination of $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ produces $\mathbf{0}$:

$$c_1 \mathbf{v}_1 + \cdots + c_k \mathbf{v}_k + (-1)\mathbf{v}_i + \mathbf{v}_j = \mathbf{0}.$$

This is impossible because $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is linearly independent. So we conclude that $T(\mathbf{v}_i) \neq T(\mathbf{v}_j)$ for all $k+1 \leq i < j \leq n$.

Finally, we show that $\mathcal{S}$ is linearly independent. Suppose

$$b_{k+1}\, T(\mathbf{v}_{k+1}) + \cdots + b_n\, T(\mathbf{v}_n) \;=\; \mathbf{0}$$

for some scalars $b_{k+1}, \dots, b_n$. We want to show that $b_{k+1} = \cdots = b_n = 0$. By linearity of T, we have

$$T(b_{k+1}\, \mathbf{v}_{k+1} + \cdots + b_n\, \mathbf{v}_n) \;=\; \mathbf{0},$$

so $b_{k+1}\mathbf{v}_{k+1} + \cdots + b_n\mathbf{v}_n \in \ker(T)$. Since $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a basis for $\ker(T)$, there are scalars $c_1, \dots, c_k$ such that

$$c_1\, \mathbf{v}_1 + \cdots + c_k\, \mathbf{v}_k \;=\; b_{k+1}\, \mathbf{v}_{k+1} + \cdots + b_n\, \mathbf{v}_n.$$

In other words, the following linear combination of $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ produces $\mathbf{0}$:

$$c_1\, \mathbf{v}_1 + \cdots + c_k\, \mathbf{v}_k + (-b_{k+1})\, \mathbf{v}_{k+1} + \cdots + (-b_n)\, \mathbf{v}_n \;=\; \mathbf{0}.$$

But $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is linearly independent, so it must be that $b_{k+1} = \cdots = b_n = 0$, as claimed. This proves that $\mathcal{S}$ is linearly independent. □