

3 Lecture 3: Jan 25

Last time

- Git

Today

- Linear algebra: vector and vector space, rank of a matrix
- Column space and Nullspace (JM Appendix A)

Notations

$$\underset{n \times 1}{\mathbf{y}} = \underset{n \times p}{\mathbf{X}} \underset{p \times 1}{\boldsymbol{\beta}} + \underset{n \times 1}{\boldsymbol{\epsilon}}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

- All vectors are column vector
- Write dimensions underneath as in $\underset{n \times p}{\mathbf{X}}$ or as $\mathbf{X} \in \mathbb{R}^{n \times p}$
- Bold upper-case letters for Matrices. Bold lower-case letters for Vectors.

Vector and vector space

(from JM Appendix A)

- A set of vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ are *linearly dependent* if there exist coefficients c_j for $j = 1, 2, \dots, n$ such that $\sum_{j=1}^n c_j \mathbf{x}_j = \mathbf{0}$ and $\|\mathbf{c}\|_2 = \sum_{j=1}^n c_j^2 > 0$. They are *linearly independent* if $\sum_{j=1}^n c_j \mathbf{x}_j = \mathbf{0}$ implies (i.e. \implies) $c_j = 0$ for all j .
- Two vectors are *orthogonal* to each other, written $\mathbf{x} \perp \mathbf{y}$, if their inner product is 0, that is $\mathbf{x}^T \mathbf{y} = \mathbf{y}^T \mathbf{x} = \sum_j x_j y_j = 0$.
- A set of vectors $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n)}$ are mutually orthogonal iff (i.e. \iff) $\mathbf{x}^{(i)T} \mathbf{x}^{(j)} = 0$ for $\forall i \neq j$.
- The most common set of vectors that are mutually orthogonal are the *elementary* vectors $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \dots, \mathbf{e}^{(n)}$, which are all zero, except for one element equal to 1, so that $\mathbf{e}_i^{(i)} = 1$ and $\mathbf{e}_j^{(i)} = 0, \forall j \neq i$.
- A *vector space* \mathcal{S} is a set of vectors that are closed under addition and scalar multiplication, that is

- if $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ are in \mathcal{S} , then $c_1\mathbf{x}^{(1)} + c_2\mathbf{x}^{(2)}$ is in \mathcal{S} .
- A vector space \mathcal{S} is *generated* or *spanned* by a set of vectors $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n)}$, written as $\mathcal{S} = \text{span}\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n)}\}$, if any vector \mathbf{x} in the vector space is a linear combination of $\mathbf{x}_i, i = 1, 2, \dots, n$.
- A set of linearly independent vectors that generate or span a space \mathcal{S} is called a *basis* of \mathcal{S} .

Example A.1

Let

$$\mathbf{x}^{(1)} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x}^{(2)} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \text{ and } \mathbf{x}^{(3)} = \begin{bmatrix} -3 \\ -1 \\ 1 \\ 3 \end{bmatrix}.$$

Then $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ are linearly independent, but $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}$, and $\mathbf{x}^{(3)}$ are linearly dependent since $5\mathbf{x}^{(1)} - 2\mathbf{x}^{(2)} + \mathbf{x}^{(3)} = \mathbf{0}$

Rank

Some matrix concepts arise from viewing columns or rows of the matrix as vectors. Assume $\mathbf{A} \in \mathbb{R}^{m \times n}$.

- $\text{rank}(\mathbf{A})$ is the maximum number of linearly independent rows or columns of a matrix.
- $\text{rank}(\mathbf{A}) \leq \min\{m, n\}$.
- A matrix is *full rank* if $\text{rank}(\mathbf{A}) = \min\{m, n\}$. It is *full row rank* if $\text{rank}(\mathbf{A}) = m$. It is *full column rank* if $\text{rank}(\mathbf{A}) = n$.
- a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is *singular* if $\text{rank}(\mathbf{A}) < n$ and *non-singular* if $\text{rank}(\mathbf{A}) = n$.
- $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}^T) = \text{rank}(\mathbf{A}^T \mathbf{A}) = \text{rank}(\mathbf{A} \mathbf{A}^T)$. (Show this in HW.)
- $\text{rank}(\mathbf{A} \mathbf{B}) \leq \min\{\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})\}$. (Hint: Columns of $\mathbf{A} \mathbf{B}$ are spanned by columns of \mathbf{A} and rows of $\mathbf{A} \mathbf{B}$ are spanned by rows of \mathbf{B} .)
- if $\mathbf{A} \mathbf{x} = \mathbf{0}_m$ for some $\mathbf{x} \neq \mathbf{0}_n$, then $\text{rank}(\mathbf{A}) \leq n - 1$.

Column space

Definition: The column space of a matrix, denoted by $\mathcal{C}(\mathbf{A})$ is the vector space spanned by the columns of the matrix, that is,

$$\mathcal{C}(\mathbf{A}) = \{\mathbf{x} : \text{there exists a vector } \mathbf{c} \text{ such that } \mathbf{x} = \mathbf{A} \mathbf{c}\}.$$

This means that if $\mathbf{x} \in \mathcal{C}(\mathbf{A})$, we can find coefficients c_j such that

$$\mathbf{x} = \sum_j c_j \mathbf{a}^{(j)}$$

where $\mathbf{a}^{(j)} = \mathbf{A}_{\cdot j}$ denotes the j^{th} column of matrix \mathbf{A} .

- The column space of a matrix consists of all vectors formed by multiplying that matrix by any vector.
- The number of basis vectors for $\mathcal{C}(\mathbf{A})$ is then the number of linearly independent columns of the matrix \mathbf{A} , and so, $\dim(\mathcal{C}(\mathbf{A})) = \text{rank}(\mathbf{A})$.
- The dimension of a space is the number of vectors in its basis.

Example A.2

Let $\mathbf{A} = \begin{bmatrix} 1 & 1 & -3 \\ 1 & 2 & -1 \\ 1 & 3 & 1 \\ 1 & 4 & 3 \end{bmatrix}$ and $\mathbf{c} = \begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix}$. Show that \mathbf{Ac} is a linear combination of columns in \mathbf{A} .

solution:

Result A.1

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

proof:

Result A.2

- (a) If $\mathbf{A} = \mathbf{BC}$, then $\mathcal{C}(\mathbf{A}) \subseteq \mathcal{C}(\mathbf{B})$.
- (b) If $\mathcal{C}(\mathbf{A}) \subseteq \mathcal{C}(\mathbf{B})$, then there exists a matrix \mathbf{C} such that $\mathbf{A} = \mathbf{BC}$.

proof:

Null space

Definition: The null space of a matrix, denoted by $\mathcal{N}(\mathbf{A})$, is $\mathcal{N}(\mathbf{A}) = \{\mathbf{y} : \mathbf{Ay} = \mathbf{0}\}$.

Result A.3

If \mathbf{A} has full-column rank, then $\mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$.

proof:

Theorem A.1

Assume $\mathbf{A} \in \mathbb{R}^{m \times n}$, then $\dim(\mathcal{C}(\mathbf{A})) = r$ and $\dim(\mathcal{N}(\mathbf{A})) = n - r$, where $r = \text{rank}(\mathbf{A})$.

See JM Appendix Theorem A.1 for the proof.

proof: Denote $\dim(\mathcal{N}(\mathbf{A}))$ by k , to be determined, and construct a set of basis vectors for $\mathcal{N}(\mathbf{A}) : \{\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \dots, \mathbf{u}^{(k)}\}$, so that $\mathbf{Au}^{(i)} = \mathbf{0}$, for $i = 1, 2, \dots, k$. Now, construct

a basis for \mathbb{R}^n by adding the vectors $\{\mathbf{u}^{(k+1)}, \dots, \mathbf{u}^{(n)}\}$, which are not in $\mathcal{N}(\mathbf{A})$. Clearly, $\mathbf{A}\mathbf{u}^{(i)} \in \mathcal{C}(\mathbf{A})$ for $i = k+1, \dots, n$, and so the span of these vectors form a subspace of $\mathcal{C}(\mathbf{A})$. These vectors $\{\mathbf{A}\mathbf{u}^{(i)}, i = k+1, \dots, n\}$ are also linearly independent from the following argument: suppose $\sum_{i=k+1}^n c_i \mathbf{A}\mathbf{u}^{(i)} = \mathbf{0}$; then $\sum_{i=k+1}^n c_i \mathbf{A}\mathbf{u}^{(i)} = \mathbf{A} [\sum_{i=k+1}^n c_i \mathbf{u}^{(i)}] = \mathbf{0}$, and hence $\sum_{i=k+1}^n c_i \mathbf{u}^{(i)}$ is a vector in $\mathcal{N}(\mathbf{A})$. Therefore, there exist b_i such that $\sum_{i=k+1}^n c_i \mathbf{u}^{(i)} = \sum_{i=1}^k b_i \mathbf{u}^{(i)}$, or $\sum_{i=1}^k b_i \mathbf{u}^{(i)} - \sum_{i=k+1}^n c_i \mathbf{u}^{(i)} = \mathbf{0}$. Since $\{\mathbf{u}^{(i)}\}$ form a basis for \mathbb{R}^n , c_i must all be zero. Therefore $\mathbf{A}\mathbf{u}^{(i)}, i = k+1, \dots, n$ are linearly independent. At this point, since $\text{span}\{\mathbf{A}\mathbf{u}^{(k+1)}, \dots, \mathbf{A}\mathbf{u}^{(n)}\} \subseteq \mathcal{C}(\mathbf{A})$, $\dim(\mathcal{C}(\mathbf{A}))$ is at least $n - k$. Suppose there is a vector \mathbf{y} that is in $\mathcal{C}(\mathbf{A})$, but not in the span; then there exists $\mathbf{u}^{(n+1)}$ so that $\mathbf{y} = \mathbf{A}\mathbf{u}^{(n+1)}$ and $\mathbf{u}^{(n+1)}$ is linearly independent of $\{\mathbf{u}^{(k+1)}, \dots, \mathbf{u}^{(n)}\}$ (and clearly not in $\mathcal{N}(\mathbf{A})$), making $n+1$ linearly independent vectors in \mathbb{R}^n . Since that is not possible, the span is equal to $\mathcal{C}(\mathbf{A})$ and $\dim(\mathcal{C}(\mathbf{A})) = n - k = r = \text{rank}(\mathbf{A})$, so that $k = \dim(\mathcal{N}(\mathbf{A})) = n - r$.

Interpretation: “dimension of column space + dimension of null space = # columns”

Mis-Interpretation: Columns space and null space are orthogonal complement to each other. They are of different orders in general! Next result gives the correct statement.