Week 2

Matrices and linear combinations (Section 2.1)

Matrix: Notation for sequence of vectors:

$$\left(\begin{bmatrix}1\\3\\5\end{bmatrix}, \begin{bmatrix}2\\4\\6\end{bmatrix}\right) \quad\to\quad \begin{bmatrix}1 & 2\\3 & 4\\5 & 6\end{bmatrix} \quad\to\quad \left(\begin{bmatrix}1 & 2\end{bmatrix}, \begin{bmatrix}3 & 4\end{bmatrix}, \begin{bmatrix}5 & 6\end{bmatrix}\right)$$

Gives another sequence of (row) vectors.

Definition 2.1: An $m \times n$ matrix is a rectangular array of real numbers with m rows and n columns.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

a_{ij}: entry in row i, column j

Dot-free notation: $A = [a_{ij}]_{i=1,j=1}^{m\ \ n}$

$\mathbb{R}^{m\times n}$: set of $m \times n$ matrices

Column notation:

$$A = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix}$$

Row notation:

$$A = \begin{bmatrix} — & \mathbf{u}_1 & — \\ — & \mathbf{u}_2 & — \\ & \vdots & \\ — & \mathbf{u}_m & — \end{bmatrix}$$

Column vector $\mathbf{v} \in \mathbb{R}^m$: $m \times 1$ matrix

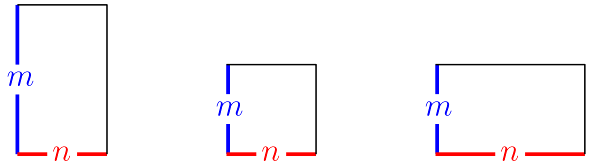
Row vector $\mathbf{u} \in \mathbb{R}^n$: $1 \times n$ matrix

Matrix addition, matrix scalar multiplication: Definition 2.2

$$\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix} + \begin{bmatrix}5 & 6\\7 & 8\end{bmatrix} = \begin{bmatrix}6 & 8\\10 & 12\end{bmatrix}, \quad 2\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix} = \begin{bmatrix}2 & 4\\6 & 8\end{bmatrix}.$$

0: zero matrix (all entries are 0)

Matrix shapes:



tall and skinny $m > n$ | square $m = n$ | short and wide $m < n$

Square matrices: Definition 2.3

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\quad\begin{bmatrix}2&0&0\\0&4&0\\0&0&5\end{bmatrix}\quad\begin{bmatrix}2&1&0\\0&4&7\\0&0&5\end{bmatrix}\quad\begin{bmatrix}2&0&0\\1&4&0\\0&7&5\end{bmatrix}\quad\begin{bmatrix}2&1&0\\1&4&7\\0&7&5\end{bmatrix}$$

identity I: $a_{ij} = \delta_{ij}$; diagonal: $j \neq i : a_{ij} = 0$; upper triangular: $j < i : a_{ij} = 0$; lower triangular: $j > i : a_{ij} = 0$; symmetric: $a_{ij} = a_{ji}$

Kronecker delta: $\delta_{ij} = 1$ if $i = j$ and 0 otherwise.

$m \times m$ identity matrix $I = [\delta_{ij}]_{i=1,j=1}^{m\ \ m}$

Matrix-vector multiplication: Notation for linear combination of the columns

$$\underbrace{7\begin{bmatrix}1\\3\\5\end{bmatrix} + 8\begin{bmatrix}2\\4\\6\end{bmatrix}}_{\text{linear combination}} = \underbrace{\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}\begin{bmatrix}7\\8\end{bmatrix}}_{\text{matrix-vector product}}.$$

Definition 2.4: Let

$$A = \begin{bmatrix}| & | & & |\\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n\\ | & | & & |\end{bmatrix} \in \mathbb{R}^{m\times n}, \quad \mathbf{x} = \begin{bmatrix}x_1\\x_2\\ \vdots\\ x_n\end{bmatrix} \in \mathbb{R}^n.$$

The vector

$$A\mathbf{x} := \sum_{j=1}^{n} x_j \mathbf{v}_j \in \mathbb{R}^m$$

is the product of A and $\mathbf{x}$.

Direct definition (without columns): Observation 2.5

$$\begin{array}{c}\mathbf{u}_1\\ \mathbf{u}_2\\ \vdots\\ \mathbf{u}_m\end{array}\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\ \vdots&\vdots&\ddots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}\begin{bmatrix}x_1\\x_2\\ \vdots\\x_n\end{bmatrix} = \begin{bmatrix}a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n\\a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n\\ \vdots\\a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n\end{bmatrix}\begin{array}{c}\mathbf{u}_1\cdot\mathbf{x}\\ \mathbf{u}_2\cdot\mathbf{x}\\ \vdots\\ \mathbf{u}_m\cdot\mathbf{x}\end{array}$$

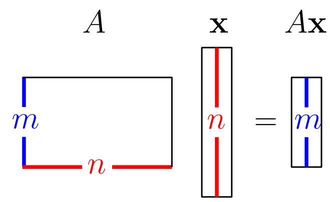
$$\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n \qquad\qquad x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n$$

Corollary 2.6: $I\mathbf{x} = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^m$.

Definition in row notation: Observation 2.7

$$A = \begin{bmatrix} — & \mathbf{u}_1 & — \\ — & \mathbf{u}_2 & — \\ & \vdots & \\ — & \mathbf{u}_m & — \end{bmatrix}, \quad A\mathbf{x} = \underbrace{\begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{x} \\ \mathbf{u}_2 \cdot \mathbf{x} \\ \vdots \\ \mathbf{u}_m \cdot \mathbf{x} \end{bmatrix}}_{\text{scalar products}}.$$

Pictorial view:



Column space and rank:

Definition 2.8: Let A be an $m \times n$ matrix. The *column space* or *image* $\mathbf{C}(A)$ of A is the span (set of all linear combinations) of the columns,

$$\mathbf{C}(A) := \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m.$$

$\mathbf{x} = \mathbf{0} \Rightarrow \mathbf{0} \in \mathbf{C}(A)$.

Fact 1.5:

$$\mathbf{C}\left(\begin{bmatrix} 2 & 3 \\ 3 & -1 \end{bmatrix}\right) = \mathbf{Span}\left(\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \end{bmatrix}\right) = \mathbb{R}^2.$$

Independent columns:

Definition 2.9: Let

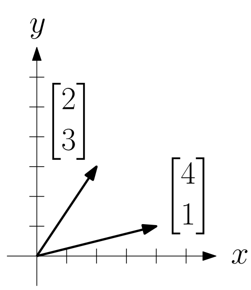
$$A = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix}.$$

Column $\mathbf{v}_j$ is *independent* if $\mathbf{v}_j$ is not a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{j-1}$. Otherwise, $\mathbf{v}_j$ is *dependent*. The *rank* of A, $\mathbf{rank}(A)$, is the number of independent columns.

$\mathbf{rank}(A) = n$: linearly independent columns

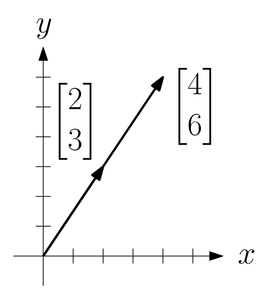
$\mathbf{rank}(A) = 0$: zero matrix

$$\mathbf{rank}\left(\begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix}\right) = 2$$



both columns are independent

$$\mathbf{rank}\left(\begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}\right) = 1$$



only first column is independent

Later: Reordering columns does not change the rank.

The independent columns span the column space:

Lemma 2.10: Let A be an $m \times n$ matrix with r independent columns, and let C be the $m \times r$ submatrix containing the independent columns. Then $\mathbf{C}(A) = \mathbf{C}(C)$.

Proof.

$\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$: the independent columns.

$\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{n-r}$: the dependent columns (order as in A)

For all j, $\mathbf{w}_j$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r, \mathbf{w}_1, \mathbf{w}_2, \ldots \mathbf{w}_{j-1}$: sequence contains all previous columns.

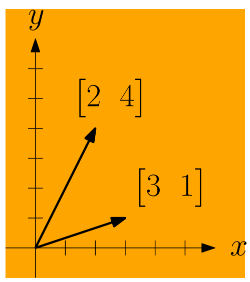
Take $\underbrace{\mathbf{Span}(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r)}_{\mathbf{C}(C)}$, add $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{n-r} \to \underbrace{\mathbf{Span}(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r, \mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{n-r})}_{\mathbf{C}(A)}$.

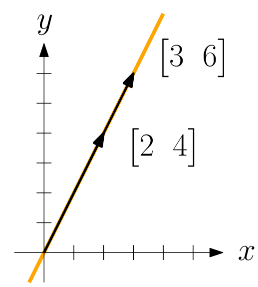
Adding linear combinations of previous vectors never changes the span (Lemma 1.23)!

□

Row space and transpose:

Row space: span of the rows

$\mathbf{R}\left(\begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix}\right)$ 

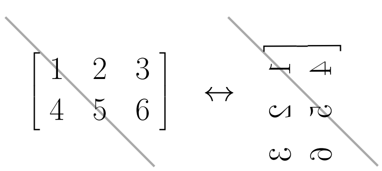
$\mathbf{R}\left(\begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}\right)$ 

Later: rank, number of independent columns = row rank, number of independent rows

To define row space, independent row, (row) rank:

- Copy& Paste from column definitions?
- Use transpose matrices!

Mirroring a matrix along the diagonal:



$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad \leftrightarrow \quad A^\top = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Definition 2.11: Let $A = [a_{ij}]_{i=1,j=1}^{m\ \ n}$ be an $m \times n$ matrix. The *transpose* of A is the $n \times m$ matrix

$$A^\top := [a_{ji}]_{i=1,j=1}^{n\ \ m}.$$

Row vector $\leftrightarrow$ column vector:

$$\begin{bmatrix}1\\3\\5\end{bmatrix}^\top = \begin{bmatrix}1 & 3 & 5\end{bmatrix}$$

$(A^\top)^\top = A$. Observation 2.12

A symmetric $\Leftrightarrow A = A^\top$.

Definition 2.13: Let A be an $m \times n$ matrix. The *row space* $\mathbf{R}(A)$ of A is the column space of the transpose,

$$\mathbf{R}(A) := \mathbf{C}(A^\top).$$

Matrix multiplication (Section 2.2)

Matrix multiplication: Notation for *several* linear combinations of the columns

Definition 2.16: Let A be an $a \times n$ matrix and

$$B = \begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_b \\ | & | & & | \end{bmatrix}$$

an $n \times b$ matrix. The $a \times b$ matrix

$$AB := \begin{bmatrix} | & | & & | \\ A\mathbf{x}_1 & A\mathbf{x}_2 & \cdots & A\mathbf{x}_b \\ | & | & & | \end{bmatrix}$$

is the product of A and B.

AB is defined exactly if number of columns of A = number of rows of B.

$\mathbf{v} = \begin{bmatrix}2\\3\end{bmatrix}, \mathbf{w} = \begin{bmatrix}3\\-1\end{bmatrix}$:

λ	μ	$\lambda\mathbf{v} + \mu\mathbf{w}$
-3	2	$\begin{bmatrix}0\\-11\end{bmatrix}$
1	-1	$\begin{bmatrix}-1\\4\end{bmatrix}$
3	0	$\begin{bmatrix}6\\9\end{bmatrix}$

$$\underbrace{\begin{bmatrix}2 & 3\\3 & -1\end{bmatrix}}_{A}\underbrace{\begin{bmatrix}-3 & 1 & 3\\2 & -1 & 0\end{bmatrix}}_{B} = \underbrace{\begin{bmatrix}0 & -1 & 6\\-11 & 4 & 9\end{bmatrix}}_{AB}$$

Direct definition: "Rows of A times columns of B" Observation 2.17

$$\underbrace{\begin{bmatrix} - & \mathbf{u}_1 & - \\ - & \mathbf{u}_2 & - \\ & \vdots & \\ - & \mathbf{u}_a & - \end{bmatrix}}_{A}\underbrace{\begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_b \\ | & | & & | \end{bmatrix}}_{B} = \underbrace{\begin{bmatrix} \mathbf{u}_1\cdot\mathbf{x}_1 & \mathbf{u}_1\cdot\mathbf{x}_2 & \cdots & \mathbf{u}_1\cdot\mathbf{x}_b \\ \mathbf{u}_2\cdot\mathbf{x}_1 & \mathbf{u}_2\cdot\mathbf{x}_2 & \cdots & \mathbf{u}_2\cdot\mathbf{x}_b \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_a\cdot\mathbf{x}_1 & \mathbf{u}_a\cdot\mathbf{x}_2 & \cdots & \mathbf{u}_a\cdot\mathbf{x}_b \end{bmatrix}}_{AB\ (ab \text{ scalar products})} = [\mathbf{u}_i \cdot \mathbf{x}_j]_{i=1,j=1}^{a\quad b}$$

$$AB = \begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix} = \begin{bmatrix}1\cdot 0+2\cdot 1 & 1\cdot 1+2\cdot 0\\3\cdot 0+4\cdot 1 & 3\cdot 1+4\cdot 0\end{bmatrix} = \begin{bmatrix}2 & 1\\4 & 3\end{bmatrix} \quad \text{("column exchange in } A\text{")}$$

$$BA = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix} = \begin{bmatrix}0\cdot 1+1\cdot 3 & 0\cdot 2+1\cdot 4\\1\cdot 1+0\cdot 3 & 1\cdot 2+0\cdot 4\end{bmatrix} = \begin{bmatrix}3 & 4\\1 & 2\end{bmatrix} \quad \text{("row exchange in } A\text{")}$$

$AB \neq BA$, matrix multiplication is not commutative.

$(AB)^\top = B^\top A^\top$ Lemma 2.19

Corollary 2.20: Let I be the $m \times m$ identity matrix. Then $IA = A$ for all $m \times n$ matrices, and $AI = A$ for all $n \times m$ matrices.

Everything is matrix multiplication:

Matrix-vector multiplication:

$$\underbrace{\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}}_{2\times 2}\underbrace{\begin{bmatrix}1\\1\end{bmatrix}}_{2\times 1} = \underbrace{\begin{bmatrix}3\\7\end{bmatrix}}_{2\times 1}.$$

Vector-matrix multiplication:

$$\underbrace{\begin{bmatrix}1 & 1\end{bmatrix}}_{1\times 2}\underbrace{\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}}_{2\times 2} = \underbrace{\begin{bmatrix}4 & 6\end{bmatrix}}_{1\times 2}.$$

Scalar product:

$$\underbrace{\begin{bmatrix}1 & 2\end{bmatrix}}_{1\times 2}\underbrace{\begin{bmatrix}3\\4\end{bmatrix}}_{2\times 1} = \underbrace{\begin{bmatrix}11\end{bmatrix}}_{1\times 1} = 11.$$

$\Rightarrow$ Another scalar product notation: $\mathbf{v}\cdot\mathbf{w} = \mathbf{v}^\top\mathbf{w}$

Outer product:

$$\underbrace{\begin{bmatrix}3\\4\end{bmatrix}}_{2\times 1}\underbrace{\begin{bmatrix}1 & 2\end{bmatrix}}_{1\times 2} = \underbrace{\begin{bmatrix}3 & 6\\4 & 8\end{bmatrix}}_{2\times 2}.$$

Lemma 2.21 Let A be an $m \times n$ matrix. The following two statements are equivalent.

(i) $\textbf{rank}(A) = 1$.

(ii) There are nonzero vectors $\mathbf{v} \in \mathbb{R}^m, \mathbf{w} \in \mathbb{R}^n$ such that A is their outer product, $A = \mathbf{v}\mathbf{w}^\top$.

Distributivity and associativity:

Lemma 2.22: Let A, B, C be three matrices Whenever the sums and products are defined, then

(i) $A(B+C) = AB + AC$ and $(A+B)C = AB + AD$ (*distributivity*);

(ii) $(AB)C = A(BC)$ (*associativity*).

Generalized associativity: brackets don't matter, also with more matrices (needs a separate proof): $(AB)(CD) = A((BC)D) = \cdots = ABCD$

CR **decomposition**:

Lemma 2.23: Let A be an $m \times n$ matrix of rank r (Definition 2.9). Let C be the $m \times r$ submatrix of A containing the independent columns. Then there exists a unique $r \times n$ matrix R such that

$$A = CR.$$

Example, $r = 1$ (we get outer product form, see Lemma 2.21):

$$\underbrace{\begin{bmatrix} 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}}_{A,2\times 3} = \underbrace{\begin{bmatrix} 2 \\ 3 \end{bmatrix}}_{C,2\times 1} \underbrace{\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}}_{R,1\times 3}.$$

Proof. A and C have the same column space (Lemma 2.10)

$\Rightarrow$ Column $\mathbf{v}_j$ of A is a linear combination of the columns of C: $\mathbf{v}_j = C\mathbf{x}_j$ ($\mathbf{x}_j \in \mathbb{R}^r$)

$$A = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix} = C \underbrace{\begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \\ | & | & & | \end{bmatrix}}_{R \in \mathbb{R}^{r\times n}} = CR.$$

C has linearly independent columns (Corollary 1.20)

$\Rightarrow$ The vectors $\mathbf{x}_j$ and hence R is unique (Lemma 1.21).

□

Example, $r = 2$:

$$A = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 2 & 4 & 1 & 4 \\ 3 & 6 & 2 & 5 \end{bmatrix}$$

	columns of A			
	$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$	$\mathbf{v}_2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$	$\mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$	$\mathbf{v}_4 = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$
$\mathbf{v}_1$	$1\mathbf{v}_1$	$2\mathbf{v}_1$		$3\mathbf{v}_1$
$\mathbf{v}_2$				
$\mathbf{v}_3$			$1\mathbf{v}_3$	$-2\mathbf{v}_3$
$\mathbf{v}_4$				
independent?	yes	no	yes	no

$$\underbrace{\begin{bmatrix} 1 & 2 & 0 & 3 \\ 2 & 4 & 1 & 4 \\ 3 & 6 & 2 & 5 \end{bmatrix}}_{A,\ 3\times 4} = \underbrace{\begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}}_{C,\ 3\times 2} \underbrace{\begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{bmatrix}}_{R,\ 2\times 4}$$