Overview

- Vectors and matrices
- Linear combinations
- Matrix operations
- Linear transformations
- Linear system, Gaussian elimination, similarity transforms, row-echelon form
- Matrix inverse, Gauss-Jordan
- Vector space, linear dependence, linear independence
- Matrix vector spaces
- Vector space basis, dimension
- Vector space sum, direct sum. Fundamental theorem of linear algebra
- Gram-Schmidt orthonormalization
- Factorizations: LU, QR

- A vector is a grouping of m scalars

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix} \in S^m, v_i \in S, \text{usually } S = \mathbb{R}.$$

- An m by n matrix is a grouping of n vectors,

$$\boldsymbol{A} = [\ \boldsymbol{a}_1 \ \ \boldsymbol{a}_2 \ \ \dots \ \ \boldsymbol{a}_n\ ] \in S^{m \times n}, \text{usually } S^{m \times n} = \mathbb{R}^{m \times n}$$

  where each vector has m scalar components $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n \in S^m$, usually $\mathbb{R}^m$.
- Matrix components with indices taking values $i \in \{1, \dots, m\}, j \in \{1, \dots, n\}$

$$\boldsymbol{A} \in \mathbb{R}^{m \times n}, \boldsymbol{A} = [a_{ij}]$$

- **Linear combination**. Let $\alpha, \beta \in S$, $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{V}$. Define a linear combination of two vectors by

$$\boldsymbol{w} = \alpha\ \boldsymbol{u} + \beta\ \boldsymbol{v} = \begin{bmatrix} \alpha\, u_1 \\ \alpha\ u_2 \\ \vdots \\ \alpha\ u_m \end{bmatrix} + \begin{bmatrix} \beta\ v_1 \\ \beta\ v_2 \\ \vdots \\ b\beta\, v_m \end{bmatrix} = \begin{bmatrix} \alpha\ u_1 + \beta\ v_1 \\ \alpha\ u_2 + \beta\ v_2 \\ \vdots \\ \alpha\, u_m + \beta\, v_m \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix}$$

- Linear combination of n vectors

$$\boldsymbol{b} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n = \begin{bmatrix} x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \cdots + x_na_{2n} \\ \vdots \\ x_1a_{m1} + x_2a_{m2} + \cdots + x_na_{mn} \end{bmatrix}$$

$$\boldsymbol{b} = x_1\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2\begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

- The sum of matrices $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{m \times n}$

$$\boldsymbol{A} = [a_{i,j}], \boldsymbol{B} = [b_{i,j}]$$

  is the matrix $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$ with components

$$\boldsymbol{C} = [c_{i,j}], c_{i,j} = a_{i,j} + b_{i,j}$$

- The scalar multiplication of matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ by $\alpha \in \mathbb{R}$ is $\boldsymbol{B} = \alpha \boldsymbol{A}$

$$\boldsymbol{A} = [a_{i,j}], \boldsymbol{B} = [b_{i,j}] = [\alpha a_{i,j}]$$

- $(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$, $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$

- Matrix-vector multiplication=linear combination ($\boldsymbol{b} \in \mathbb{R}^m$, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{x} \in \mathbb{R}^n$)

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n = \begin{bmatrix} x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \cdots + x_na_{2n} \\ \vdots \\ x_1a_{m1} + x_2a_{m2} + \cdots + x_na_{mn} \end{bmatrix}$$

- Matrix-matrix multiplications $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{X}$ ($\boldsymbol{B} \in \mathbb{R}^{m \times p}$, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{X} \in \mathbb{R}^{n \times p}$)

$$\boldsymbol{B} = [\ \boldsymbol{b}_1 \ \ \ldots \ \ \boldsymbol{b}_p\ ] = \boldsymbol{A}[\ \boldsymbol{x}_1 \ \ \ldots \ \ \boldsymbol{x}_p\ ] = [\ \boldsymbol{A}\boldsymbol{x}_1 \ \ \ldots \ \ \boldsymbol{A}\boldsymbol{x}_p\ ]$$

- Vectors are single-column matrices, $\boldsymbol{b} \in \mathbb{R}^m$ is shorthand for $\boldsymbol{b} \in \mathbb{R}^{m \times 1}$
- Matrix multiplication is associative

$$(\boldsymbol{A}\boldsymbol{B})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{B}\boldsymbol{C})$$

- Matrix multiplication is not commutative, i.e., there do exist $\boldsymbol{A}, \boldsymbol{B}$ such that

$$\boldsymbol{A}\boldsymbol{B} \neq \boldsymbol{B}\boldsymbol{A}$$

- Grouping of m scalars into a column vector $\boldsymbol{v}$ is an arbitrary choice

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix} \in S^{m \times 1}, v_i \in S, \text{usually } S = \mathbb{R}.$$

- Introduce transposition to switch between the two groupings

$$\boldsymbol{v}^T = [\ v_1 \quad v_2 \quad \dots \quad v_m\ ] \in S^{1 \times m}, v_i \in S, \text{usually } S = \mathbb{R}.$$

- A preferred type of grouping is useful in calculations
- Preferred grouping: column vectors such that $\boldsymbol{v} \in S^m$ is understood to signify a column vector. When explicitly required, write $\boldsymbol{v} \in S^{m \times 1}$
- Linear combination of row vectors

$$\boldsymbol{b}^T = x_1 \boldsymbol{a}_1^T + x_2 \boldsymbol{a}_2^T + \dots + x_n \boldsymbol{a}_n^T$$

- Matrix-vector multiplication has been introduced as

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n = \begin{bmatrix} x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \cdots + x_na_{2n} \\ \vdots \\ x_1a_{m1} + x_2a_{m2} + \cdots + x_na_{mn} \end{bmatrix}$$

- Matrix-vector multiplication can be seen a "rows over columns rule"

$$\boldsymbol{A}\boldsymbol{x} = [\ \boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \ldots \quad \boldsymbol{a}_n\ ]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n$$

- "Rows over columns" also works for components

$$\boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \cdots + x_na_{2n} \\ \vdots \\ x_1a_{m1} + x_2a_{m2} + \cdots + x_na_{mn} \end{bmatrix}$$

- Scalar product: $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^m$, $\boldsymbol{u}^T\boldsymbol{v} = \boldsymbol{u}\cdot\boldsymbol{v} = u_1v_1 + u_2v_2 + \cdots + u_mv_m$
- 2-norm: $\boldsymbol{u} \in \mathbb{R}^m$, $\|\boldsymbol{u}\| = (\boldsymbol{u}^T\boldsymbol{u})^{1/2}$
- Angle between two vectors:

$$\cos(\theta) = \frac{\boldsymbol{v}^T\boldsymbol{w}}{\|\boldsymbol{v}\|\,\|\boldsymbol{w}\|}$$

- Transpose: swap between rows and columns $\boldsymbol{A} = [\ \boldsymbol{a}_1\ \ \ldots\ \ \boldsymbol{a}_n\ ]$

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \boldsymbol{A}^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1^T \\ \vdots \\ \boldsymbol{a}_n^T \end{bmatrix}$$

- Linear combination of rows as vector-matrix product

$$\boldsymbol{b}^T = x_1\boldsymbol{a}_1^T + x_2\boldsymbol{a}_2^T + \cdots + x_n\boldsymbol{a}_n^T = \boldsymbol{x}^T\boldsymbol{A}^T$$

- Transposition of product (multiple linear combinations): $(\boldsymbol{A}\boldsymbol{B})^T = \boldsymbol{B}^T\boldsymbol{A}^T$

- Matrices often exhibit some intrinsic structure, for example

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{C} \\ \boldsymbol{C} & \boldsymbol{B} \end{bmatrix}, \boldsymbol{D} = \begin{bmatrix} \boldsymbol{E} & \boldsymbol{F} \\ \boldsymbol{F} & \boldsymbol{E} \end{bmatrix}$$

  $\boldsymbol{A}, \boldsymbol{D} \in \mathbb{R}^{2m \times 2n}$, $\boldsymbol{B}, \boldsymbol{C}, \boldsymbol{E}, \boldsymbol{F} \in \mathbb{R}^{m \times n}$ (other structures possible)

- Addition over compatible block dimensions

$$\boldsymbol{A} + \boldsymbol{D} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{C} \\ \boldsymbol{C} & \boldsymbol{B} \end{bmatrix} + \begin{bmatrix} \boldsymbol{E} & \boldsymbol{F} \\ \boldsymbol{F} & \boldsymbol{E} \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}+\boldsymbol{E} & \boldsymbol{C}+\boldsymbol{F} \\ \boldsymbol{C}+\boldsymbol{F} & \boldsymbol{B}+\boldsymbol{E} \end{bmatrix}$$

- Multiplication, "row over columns" for compatible block dimensions

$$\boldsymbol{A}\boldsymbol{D} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{C} \\ \boldsymbol{C} & \boldsymbol{B} \end{bmatrix} \begin{bmatrix} \boldsymbol{E} & \boldsymbol{F} \\ \boldsymbol{F} & \boldsymbol{E} \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}\boldsymbol{E}+\boldsymbol{C}\boldsymbol{F} & \boldsymbol{B}\boldsymbol{F}+\boldsymbol{C}\boldsymbol{E} \\ \boldsymbol{C}\boldsymbol{E}+\boldsymbol{B}\boldsymbol{F} & \boldsymbol{C}\boldsymbol{F}+\boldsymbol{B}\boldsymbol{E} \end{bmatrix}$$

- Matrix block transposition

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{U} & \boldsymbol{V} \\ \boldsymbol{X} & \boldsymbol{Y} \end{bmatrix}, \boldsymbol{M}^T = \begin{bmatrix} \boldsymbol{U}^T & \boldsymbol{X}^T \\ \boldsymbol{V}^T & \boldsymbol{Y}^T \end{bmatrix}$$

- $T\colon \mathbb{R}^n \to \mathbb{R}^m$, mapping of vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$
- Of special interest: linear mappings that preserve linear combinations

$$T(\alpha \boldsymbol{u} + \beta \boldsymbol{v}) = \alpha T(\boldsymbol{u}) + \beta T(\boldsymbol{v})$$

- Matrix $\boldsymbol{B}$ of linear transformation

$$T(\boldsymbol{v}) = T(v_1 \boldsymbol{e}_1 + \cdots + v_n \boldsymbol{e}_n) = v_1 T(\boldsymbol{e}_1) + \cdots + v_n T(\boldsymbol{e}_n)$$

$$\boldsymbol{B} = [\ T(\boldsymbol{e}_1) \quad T(\boldsymbol{e}_2) \quad \ldots \quad T(\boldsymbol{e}_n)\ ] \in \mathbb{R}^{m \times n}$$

$$T(\boldsymbol{v}) = \boldsymbol{B}\boldsymbol{v}$$

- Consider two successive linear mappings $S\colon \mathbb{R}^p \to \mathbb{R}^n$, $T\colon \mathbb{R}^n \to \mathbb{R}^m$

$$\boldsymbol{v} = S(\boldsymbol{u}) = \boldsymbol{A}\boldsymbol{u}, \boldsymbol{w} = T(\boldsymbol{v}) = \boldsymbol{B}\boldsymbol{v}$$

- Linear mapping composition, $U\colon \mathbb{R}^p \to \mathbb{R}^m$, $U = T \circ S$

$$\boldsymbol{w} = \boldsymbol{C}\boldsymbol{u} = U(\boldsymbol{u}) = T(S(\boldsymbol{u})) = T(\boldsymbol{A}\boldsymbol{u}) = \boldsymbol{B}\boldsymbol{A}\boldsymbol{u} \Rightarrow \boldsymbol{C} = \boldsymbol{B}\boldsymbol{A}$$

Matrix of composition is matrix product of individual mappings.

- Stretching, $T\colon \mathbb{R}^m \to \mathbb{R}^m$, $T(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v}$

$$\boldsymbol{A} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_m \end{bmatrix}$$

- Orthogonal projection of $\boldsymbol{v} \in \mathbb{R}^m$ along $\boldsymbol{u} \in \mathbb{R}^m$, $\|\boldsymbol{u}\| = 1$, $T(\boldsymbol{v}) = \boldsymbol{P_u}\boldsymbol{v}$

$$\boldsymbol{P_u} = \boldsymbol{u}\boldsymbol{u}^T$$

- Reflection across vector $\boldsymbol{w} \in \mathbb{R}^m$, $\|\boldsymbol{w}\| = 1$, $T(\boldsymbol{v}) = \boldsymbol{R}\boldsymbol{v}$

$$\boldsymbol{R} = 2\boldsymbol{w}\boldsymbol{w}^T - \boldsymbol{I}$$

- Rotation in $\mathbb{R}^2$

$$\boldsymbol{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

- Component form

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

- Matrix form $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^m$
- Simple systems:
  - $\boldsymbol{A}$ diagonal
  - $\boldsymbol{A}$ triangular
- Solve $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ by bringing to simpler form. Gaussian elimination: triangular

$$\begin{cases} x_1 + 2x_2 - x_3 & = & 2 \\ 2x_1 - x_2 + x_3 & = & 2 \\ 3x_1 - x_2 - x_3 & = & 1 \end{cases}, \begin{cases} x_1 + 2x_2 - x_3 & = & 2 \\ -5x_2 + 3x_3 & = & -2 \\ -7x_2 + 2x_3 & = & -5 \end{cases}, \begin{cases} x_1 + 2x_2 - x_3 & = & 2 \\ -5x_2 + 3x_3 & = & -2 \\ -\frac{11}{5}x_3 & = & -\frac{11}{5} \end{cases}$$

- Work with bordered matrix

$$\left[\begin{array}{cccc} 1 & 2 & -1 & 2 \\ 2 & -1 & 1 & 2 \\ 3 & -1 & -1 & 1 \end{array}\right] \sim \left[\begin{array}{cccc} 1 & 2 & -1 & 2 \\ 0 & -5 & 3 & -2 \\ 0 & -7 & 2 & -5 \end{array}\right] \sim \left[\begin{array}{cccc} 1 & 2 & -1 & 2 \\ 0 & -5 & 3 & -2 \\ 0 & 0 & -\frac{11}{5} & -\frac{11}{5} \end{array}\right]$$

- Can obtain no, unique, infinite solutions

$$\left\{\begin{array}{rl} x_1+2x_2+3x_3 & =3 \\ x_2+x_3 & =1 \\ 0 & =0 \end{array}\right. \text{Infinite}, \left\{\begin{array}{rl} x_1+2x_2+3x_3 & =3 \\ x_2+x_3 & =1 \\ 0 & =1 \end{array}\right. \text{None}$$

- Analyze by bring to reduced row echelon form
  - All zero rows are below non-zero rows
  - First non-zero entry on a row is called the *leading entry*
  - In each non-zero row, the leading entry is to the left of lower leading entries
  - Each leading entry equals 1 and is the only non-zero entry in its column

- Step k in Gaussian elimination can be seen as multiplication with

$$\boldsymbol{L}_k = \begin{pmatrix} 1 & \dots & 0 & \dots & 1 \\ 0 & \ddots & 0 & \dots & 0 \\ 0 & \dots & 1 & \dots & 0 \\ 0 & \dots & -l_{k+1,k} & \dots & 0 \\ 0 & \dots & -l_{k+2,k} & \dots & 0 \\ \vdots & \dots & \vdots & \ddots & \vdots \\ 0 & \dots & -l_{m,k} & \dots & 1 \end{pmatrix}, \boldsymbol{L}_k^{-1} = \begin{pmatrix} 1 & \dots & 0 & \dots & 1 \\ 0 & \ddots & 0 & \dots & 0 \\ 0 & \dots & 1 & \dots & 0 \\ 0 & \dots & l_{k+1,k} & \dots & 0 \\ 0 & \dots & l_{k+2,k} & \dots & 0 \\ \vdots & \dots & \vdots & \ddots & \vdots \\ 0 & \dots & l_{m,k} & \dots & 1 \end{pmatrix}$$

- $\boldsymbol{A} \in \mathbb{R}^{m \times m}$ is invertible if there exists $\boldsymbol{X} \in \mathbb{R}^{m \times m}$ such that

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{X}\boldsymbol{A} = \boldsymbol{I}$$

- Notation $\boldsymbol{X} = \boldsymbol{A}^{-1}$, is the *inverse* of $\boldsymbol{A}$.

- When does a matrix inverse exist? $\boldsymbol{A} \in \mathbb{R}^{m\times m}$
  a $\boldsymbol{A}$ invertible
  b $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ has a unique solution for all $\boldsymbol{b} \in \mathbb{R}^m$
  c $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ has a unique solution
  d The reduced row echelon form of $\boldsymbol{A}$ is $\boldsymbol{I}$
  e $\boldsymbol{A}$ can be written as product of elementary matrices

$$a \Rightarrow b \Rightarrow c \Rightarrow d \Rightarrow e \Rightarrow a$$

- $\boldsymbol{X}$ is inverse if $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{I}$ or

$$\boldsymbol{A}[\ \boldsymbol{x}_1\ \ \boldsymbol{x}_2\ \ \ldots\ \ \boldsymbol{x}_m\ ] = [\ \boldsymbol{A}\boldsymbol{x}_1\ \ \boldsymbol{A}\boldsymbol{x}_2\ \ \ldots\ \ \boldsymbol{A}\boldsymbol{x}_m\ ] = [\ \boldsymbol{e}_1\ \ \boldsymbol{e}_2\ \ \ldots\ \ \boldsymbol{e}_m\ ]$$

- Gauss-Jordan: similar to Gauss elimination

$$[\ \boldsymbol{A}\ \ |\ \ \boldsymbol{I}\ ] \sim [\ \boldsymbol{I}\ \ |\ \ \boldsymbol{X}\ ]$$

- $(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$, $(\boldsymbol{A}^T)^{-1} = (\boldsymbol{A}^{-1})^T$

Addition rules for	$\forall \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in V$
$\boldsymbol{a}+\boldsymbol{b} \in V$	Closure
$\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})=(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}$	Associativity
$\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$	Commutativity
$\boldsymbol{0}+\boldsymbol{a}=\boldsymbol{a}$	Zero vector
$\boldsymbol{a}+(-\boldsymbol{a})=\boldsymbol{0}$	Additive inverse
Scaling rules for	$\forall \boldsymbol{a}, \boldsymbol{b} \in V, \forall x, y \in S$
$x\boldsymbol{a} \in V$	Closure
$x(\boldsymbol{a}+\boldsymbol{b})=x\boldsymbol{a}+x\boldsymbol{b}$	Distributivity
$(x+y)\boldsymbol{a}=x\boldsymbol{a}+y\boldsymbol{a}$	Distributivity
$x(y\boldsymbol{a})=(xy)\boldsymbol{a}$	Composition
$1\boldsymbol{a}=\boldsymbol{a}$	Scalar identity

- The vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \in \mathcal{V}$, are linearly dependent if there exist n scalars, $x_1, \ldots, x_n \in \mathcal{S}$, at least one of which is different from zero such that

$$x_1\boldsymbol{a}_1 + \ldots x_n\boldsymbol{a}_n = \boldsymbol{0}$$

- The vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \in \mathcal{V}$, are linearly independent if the only n scalars, $x_1, \ldots, x_n \in \mathcal{S}$, that satisfy

$$x_1\boldsymbol{a}_1 + \ldots x_n\boldsymbol{a}_n = \boldsymbol{0}, \text{are } x_1 = x_2 = \cdots = x_n = 0$$

- A set of vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in \mathcal{V}$ is a basis for vector space $\mathcal{V}$ if:
  1. $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ are linearly independent;
  2. $\operatorname{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\} = \mathcal{V}$.
- The number of vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in \mathcal{V}$ within a basis is the dimension of the vector space $\mathcal{V}$.
- The column space (or range) of matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is the set of vectors reachable by linear combination of the matrix column vectors

$$C(\boldsymbol{A}) = \operatorname{range}(\boldsymbol{A}) = \{\boldsymbol{b} \in \mathbb{R}^m | \exists \boldsymbol{x} \in \mathbb{R}^n \text{ such that } \boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}\} \subseteq \mathbb{R}^m$$

- *Left null space*, $N(\boldsymbol{A}^T) = \{\boldsymbol{y} \in \mathbb{R}^m | \boldsymbol{A}^T \boldsymbol{y} = 0\} \subseteq \mathbb{R}^m$, the part of $\mathbb{R}^m$ *not reachable* by linear combination of columns of $\boldsymbol{A}$
- The null space of a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is the set

$$N(\boldsymbol{A}) = \operatorname{null}(\boldsymbol{A}) = \{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\} \subseteq \mathbb{R}^n$$

- The *row space* of $\boldsymbol{A}$ as

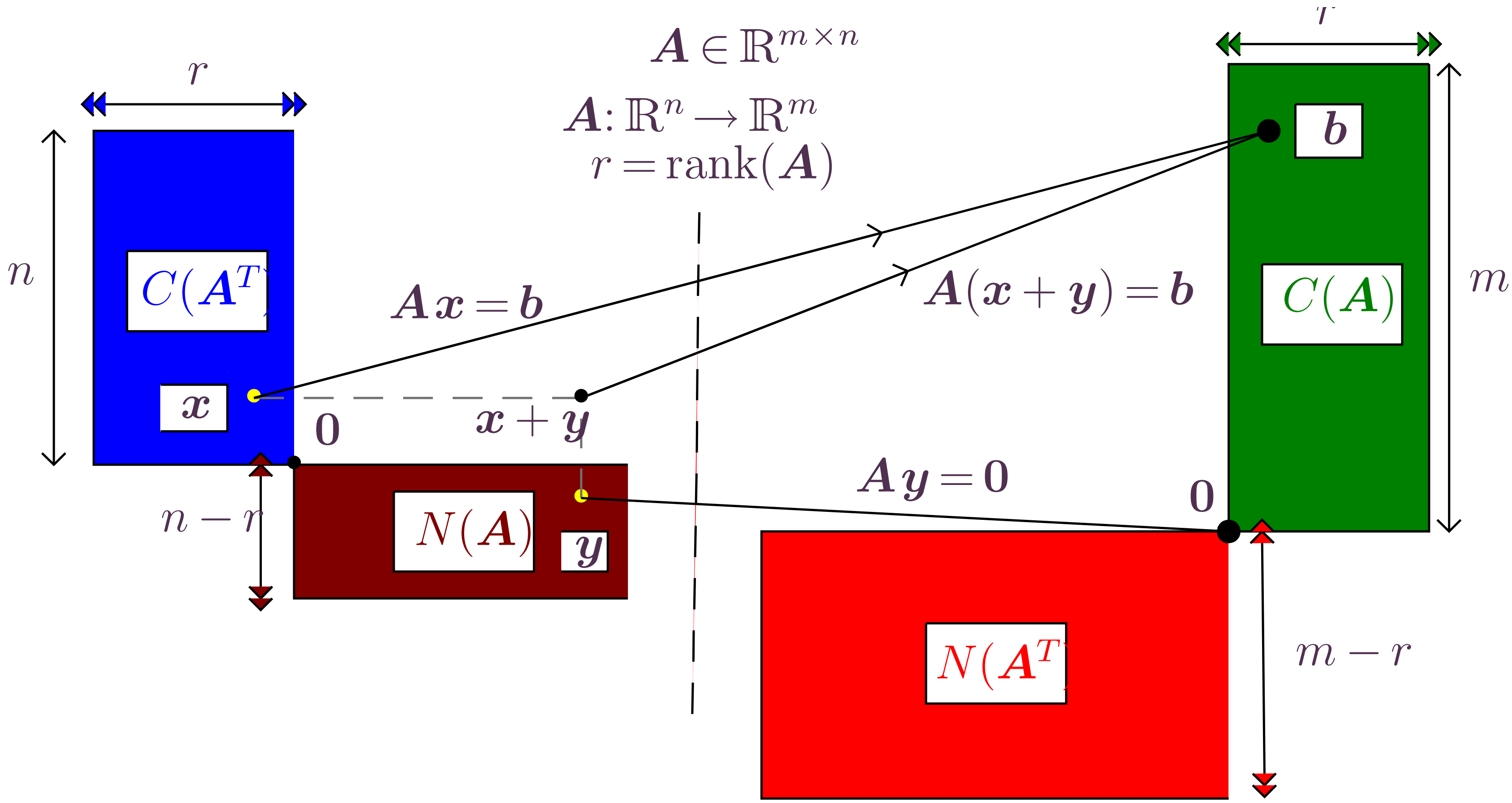
$$R(\boldsymbol{A}) = C(\boldsymbol{A}^T) = \{\boldsymbol{c} \in \mathbb{R}^n | \exists \boldsymbol{y} \in \mathbb{R}^m \text{ such that } \boldsymbol{c} = \boldsymbol{A}^T \boldsymbol{y}\} \subseteq \mathbb{R}^n$$

- Given two vector subspaces $(\mathcal{U},\mathcal{S},+)$, $(\mathcal{V},\mathcal{S},+)$ of the space $(\mathcal{W},\mathcal{S},+)$, the sum is the set $\mathcal{U}+\mathcal{V}=\{\boldsymbol{u}+\boldsymbol{v}|\ \boldsymbol{u}\in\mathcal{U},\boldsymbol{v}\in\mathcal{V}\}$.
- Given two vector subspaces $(\mathcal{U},\mathcal{S},+)$, $(\mathcal{V},\mathcal{S},+)$ of the space $(\mathcal{W},\mathcal{S},+)$, the direct sum is the set $\mathcal{U}\oplus\mathcal{V}=\{\boldsymbol{u}+\boldsymbol{v}|\ \exists!\boldsymbol{u}\in\mathcal{U},\exists!\boldsymbol{v}\in\mathcal{V}\}$. (unique decomposition)
- Given two vector subspaces $(\mathcal{U},\mathcal{S},+)$, $(\mathcal{V},\mathcal{S},+)$ of the space $(\mathcal{W},\mathcal{S},+)$, the intersection is the set

$$\mathcal{U}\cap\mathcal{V}=\{\boldsymbol{x}|\boldsymbol{x}\in\mathcal{U},\boldsymbol{x}\in\mathcal{V}\}.$$

- Two vector subspaces $(\mathcal{U},\mathcal{S},+)$, $(\mathcal{V},\mathcal{S},+)$ of the space $(\mathcal{W},\mathcal{S},+)$ are orthogonal subspaces, denoted $\mathcal{U}\perp\mathcal{V}$ if $\boldsymbol{u}^T\boldsymbol{v}=0$ for any $\boldsymbol{u}\in\mathcal{U},\boldsymbol{v}\in\mathcal{V}$.
- Two vector subspaces $(\mathcal{U},\mathcal{S},+)$, $(\mathcal{V},\mathcal{S},+)$ of the space $(\mathcal{W},\mathcal{S},+)$ are orthogonal complements, denoted $\mathcal{U}=\mathcal{V}^\perp$, $\mathcal{V}=\mathcal{U}^\perp$ if $\mathcal{U}\perp\mathcal{V}$ and $\mathcal{U}+\mathcal{V}=\mathcal{W}$.
- Orthogonal complement subspaces form a direct sum $\mathcal{U}=\mathcal{V}^\perp$, $\mathcal{V}=\mathcal{U}^\perp\Rightarrow$

$$\mathcal{U}+\mathcal{V}=\mathcal{U}\oplus\mathcal{V}$$



$\mathbb{R}^n = C(\boldsymbol{A}^T) \oplus N(\boldsymbol{A})$ usually: $m \geqslant n$ $\mathbb{R}^m = N(\boldsymbol{A}^T) \oplus C(\boldsymbol{A})$

$C(\boldsymbol{A}^T) \perp N(\boldsymbol{A})$ $N(\boldsymbol{A}^T) \perp C(\boldsymbol{A})$

- $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$, $\boldsymbol{Q}$ orthogonal, $\boldsymbol{R}$ triangular

$$\boldsymbol{A} = [\ \boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \ldots\ \ \boldsymbol{a}_n\ ] = [\ \boldsymbol{q}_1\ \ \boldsymbol{q}_2\ \ \ldots\ \ \boldsymbol{q}_n\ ] \begin{bmatrix} r_{11} & r_{12} & \ldots & r_{1n} \\ 0 & r_{22} & \ldots & r_{2n} \\ \vdots & & \ddots & \\ 0 & 0 & \ldots & r_{nn} \end{bmatrix} = \boldsymbol{Q}\boldsymbol{R}$$

- Identify on both sides to obtain

$$\begin{aligned} \boldsymbol{a}_1 &= r_{11}\,\boldsymbol{q}_1 \\ \boldsymbol{a}_2 &= r_{12}\,\boldsymbol{q}_1 + r_{22}\,\boldsymbol{q}_2 \\ \boldsymbol{a}_3 &= r_{13}\,\boldsymbol{q}_1 + r_{23}\,\boldsymbol{q}_2 + r_{33}\,\boldsymbol{q}_3 \\ &\vdots \\ \boldsymbol{a}_n &= r_{1n}\,\boldsymbol{q}_1 + r_{2n}\,\boldsymbol{q}_2 + r_{3n}\,\boldsymbol{q}_3 + \cdots + r_{nn}\boldsymbol{q}_n \end{aligned}$$

$$\begin{aligned} \boldsymbol{q}_1 &= \boldsymbol{a}_1 / r_{11} \\ \boldsymbol{q}_2 &= (\boldsymbol{a}_2 - r_{12}\,\boldsymbol{q}_1) / r_{22} \\ \boldsymbol{q}_3 &= (\boldsymbol{a}_3 - r_{13}\,\boldsymbol{q}_1 - r_{23}\,\boldsymbol{q}_2) / r_{33} \\ &\vdots \end{aligned}$$

- Gaussian elimination produces a sequence matrices similar to $\boldsymbol{A}\in\mathbb{R}^{m\times m}$

$$\boldsymbol{A}=\boldsymbol{A}^{(0)}\sim\boldsymbol{A}^{(1)}\sim\cdots\sim\boldsymbol{A}^{(k)}\sim\cdots\sim\boldsymbol{A}^{(m-1)}$$

- Step k produces zeros underneath diagonal position (k,k)
- Step k can be represented as multiplication by matrix

$$\boldsymbol{A}^{(k)}=\boldsymbol{L}_k\boldsymbol{A}^{(k-1)},\ \boldsymbol{L}_k=\begin{pmatrix} 1 & \dots & 0 & \dots & 1 \\ 0 & \ddots & 0 & \dots & 0 \\ 0 & \dots & 1 & \dots & 0 \\ 0 & \dots & -l_{k+1,k} & \dots & 0 \\ \vdots & \dots & \vdots & \ddots & \vdots \\ 0 & \dots & -l_{m,k} & \dots & 1 \end{pmatrix},\ l_{j,k}=\frac{a_{j,k}^{(k-1)}}{a_{k,k}^{(k-1)}},\ \boldsymbol{A}^{(k)}=[a_{i,j}^{(k)}]$$

- All $m-1$ steps produce an upper triangular matrix

$$\boldsymbol{L}_{m-1}\dots\boldsymbol{L}_2\boldsymbol{L}_1\boldsymbol{A}=\boldsymbol{U}\Rightarrow\boldsymbol{A}=\boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1}\dots\boldsymbol{L}_{m-1}^{-1}\boldsymbol{U}=\boldsymbol{L}\boldsymbol{U}$$

- With permutations $\boldsymbol{PA}=\boldsymbol{LU}$ (Matlab `[L,U,P]=lu(A), A=P'*L*U`)

- With known LU-factorization: $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}\Rightarrow(\boldsymbol{L}\boldsymbol{U})\boldsymbol{x}=\boldsymbol{P}\boldsymbol{b}\Rightarrow\boldsymbol{L}(\boldsymbol{U}\boldsymbol{x})=\boldsymbol{P}\boldsymbol{b}$
- To solve $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$:
  1. Carry out LU-factorization: $\boldsymbol{P}^T\boldsymbol{L}\boldsymbol{U}=\boldsymbol{A}$
  2. Solve $\boldsymbol{L}\boldsymbol{y}=\boldsymbol{c}=\boldsymbol{P}\boldsymbol{b}$ by forward substitution to find $\boldsymbol{y}$
  3. Solve $\boldsymbol{U}\boldsymbol{x}=\boldsymbol{y}$ by backward substitution
- FLOP = floating point operation = one multiplication and one addition
- Operation counts: how many FLOPS in each step?
  1. Each $\boldsymbol{L}_k\boldsymbol{A}^{(k-1)}$ costs $(m-k)^2$ FLOPS. Overall

$$(m-1)^2+(m-2)^2+\cdots+1^2=\frac{m(m-1)(2m-1)}{6}\approx\frac{m^3}{3}$$

  2. Forward substitution step k costs k flops

$$1+2+\cdots+m=\frac{m(m+1)}{2}\approx\frac{m^2}{2}$$

  3. Backward substitution cost is identical $m(m+1)/2\approx m^2/2$/

- Orthonormalization of columns of $\boldsymbol{A}$ is also a factorization

$$\boldsymbol{A} = [\ \boldsymbol{a}_1 \ \ \boldsymbol{a}_2 \ \ \ldots \ \ \boldsymbol{a}_n\ ] = [\ \boldsymbol{q}_1 \ \ \boldsymbol{q}_2 \ \ \ldots \ \ \boldsymbol{q}_n\ ] \begin{bmatrix} r_{11} & r_{12} & \ldots & r_{1n} \\ 0 & r_{22} & \ldots & r_{2n} \\ \vdots & & \ddots & \\ 0 & 0 & \ldots & r_{nn} \end{bmatrix} = \boldsymbol{Q}\boldsymbol{R}$$

$$\begin{aligned} \boldsymbol{a}_1 &= r_{11}\,\boldsymbol{q}_1 \\ \boldsymbol{a}_2 &= r_{12}\,\boldsymbol{q}_1 + r_{22}\,\boldsymbol{q}_2 \\ \boldsymbol{a}_3 &= r_{13}\,\boldsymbol{q}_1 + r_{23}\,\boldsymbol{q}_2 + r_{33}\,\boldsymbol{q}_3 \\ &\vdots \\ \boldsymbol{a}_n &= r_{1n}\,\boldsymbol{q}_1 + r_{2n}\,\boldsymbol{q}_2 + r_{3n}\,\boldsymbol{q}_3 + \cdots + r_{nn}\boldsymbol{q}_n \end{aligned}$$

$$\begin{aligned} \boldsymbol{q}_1 &= \boldsymbol{a}_1 / r_{11} \\ \boldsymbol{q}_2 &= (\boldsymbol{a}_2 - r_{12}\,\boldsymbol{q}_1) / r_{22} \\ \boldsymbol{q}_3 &= (\boldsymbol{a}_3 - r_{13}\,\boldsymbol{q}_1 - r_{23}\,\boldsymbol{q}_2) / r_{33} \\ &\vdots \end{aligned}$$

- Operation count:
  - $r_{jk} = \boldsymbol{q}_j^T \boldsymbol{a}_k$ costs m FLOPS
  - There are $1 + 2 + \cdots + n$ components in $\boldsymbol{R}$, Overall cost $n(n+1)m/2$
- With permutations $\boldsymbol{A}\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{R}$ (Matlab `[Q,R,P]=qr(A)`)

- With known QR-factorization: $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}\Rightarrow(\boldsymbol{Q}\boldsymbol{R}\boldsymbol{P}^T)\boldsymbol{x}=\boldsymbol{b}\Rightarrow\boldsymbol{R}\boldsymbol{y}=\boldsymbol{Q}^T\boldsymbol{b}$
- To solve $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$:
  1. Carry out QR-factorization: $\boldsymbol{Q}\boldsymbol{R}\boldsymbol{P}^T=\boldsymbol{A}$
  2. Compute $\boldsymbol{c}=\boldsymbol{Q}^T\boldsymbol{b}$
  3. Solve $\boldsymbol{R}\boldsymbol{y}=\boldsymbol{c}$ by backward substitution
  4. Find $\boldsymbol{x}=\boldsymbol{P}^T\boldsymbol{y}$
- Operation counts: how many FLOPS in each step?
  1. QR-factorization $m^2(m+1)/2\approx m^3/2$
  2. Compute $\boldsymbol{c}$, m^2
  3. Backward substitution $m(m+1)/2\approx m^2/2$