Math Foundations

Meng Song

Jan 7, 2022

Matrix Form of linear equations

- $\begin{cases} x_1 - 2x_2 = 1 \\ 3x_1 + 2x_2 = 11 \end{cases}$
  - Two equations
  - Two unknowns/variables
- Linear equations can be written in the matrix form
  - $\begin{bmatrix} 1 & -2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix}$
  - Ax=b
    - A: m*n matrix
    - x: n*1 vector
    - b: m*1 vector

Gaussian Elimination (Row reduction)

$$A = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix} \xrightarrow{R_2 \leftarrow R_2 - 2R_1} \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 3 & 6 & 8 & 10 \end{bmatrix} \xrightarrow{R_3 \leftarrow R_3 - 3R_1} \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 2 & 4 \end{bmatrix}$$

$$\xrightarrow{R_3 \leftarrow R_3 - R_2} \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U$$

pivot free pivot free

- **Pivots**: the leftmost non-zero element on each row
- $A \to U$

Row reduction

- $\begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_2 \leftarrow R_2/2} \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$
- $A \rightarrow U \rightarrow R$
- *Reduced echelon form (R)*
  - All pivots= 1
  - Zeros below and left to pivots

$$\begin{pmatrix} \boxed{\star} & \star & \star & \star & \star \\ 0 & \boxed{\star} & \star & \star & \star \\ 0 & 0 & 0 & \boxed{\star} & \star \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$\star$ = any number

$\boxed{\star}$ = any nonzero number

Gaussian Elimination (Row reduction)

- Elimination operations:
  - Swap two rows
  - Subtract a scalar multiple of one row from another
  - Multiply a row by a non-zero scalar
- R= $E_n E_{n-1} \ldots E_1 \mathrm{A}$
- $\begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix}$

R **E4** **E3** **E2** **E1** **A**

Range

- Range of A is the **linear combination** of the column vectors that contain the pivots.
- $\begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$
- range(A) = $c_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + c_3 \begin{bmatrix} 2 \\ 6 \\ 8 \end{bmatrix}, \forall c_1, c_3 \in \mathbb{R}$
- range = column space

Rank

- $\begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

  A **R**

- *rank(A) = # of pivots in R*
- *rank(A) = # of linearly independent columns = dimension of range*
- *rank(A) = # of nonzero rows in R*
- *rank(A)* $\leq$ *min(m, n)*
  - If rank(A)=m, $Ax = b$ has solution for $\forall b \in \mathbb{R}$

Nullspace

- The nullspace N(A) in $\mathbb{R}^n$ contains all solutions x to Ax = **0**, including x= **0**.
- $Ax = \mathbf{0} \rightarrow Rx = \mathbf{0}$
  - Row reduction will not change the solutions
  - $\begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$
  - $\begin{cases} x_1 + 2x_2 + 2x_3 + 2x_4 = 0 \\ x_3 + 2x_4 = 0 \end{cases}$

Nullspace

- Identify the free variables
  - $$\begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

    ⬆ (column 2) ⬆ (column 4)

    free columns

  - Free columns → free variables
  - Pivot columns → pivot variables

Nullspace

- Represent the pivot variables with respect to the free variables
- $\begin{cases} x_1 + 2x_2 + 2x_3 + 2x_4 = 0 \\ x_3 + 2x_4 = 0 \end{cases} \rightarrow \begin{cases} x_1 = -2x_2 + 2x_4 \\ x_2 = x_2 \\ x_3 = -2x_4 \\ x_4 = x_4 \end{cases}$
- Writing it in matrix form gives us the **nullspace** of A
  - N(A): $x = \begin{bmatrix} -2 & 2 \\ 1 & 0 \\ 0 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ x_4 \end{bmatrix}$, where $\forall x_2 , x_4 \in \mathbb{R}$
- dimension of N(A) = n – rank(A)

Solve Ax=b

- Row reduce the augmented matrix $[A \quad b]$ to $[R \quad d]$
  - Row reduction will not change the solutions
  - If $d_m = 0$, the equations are solvable, otherwise no solution
- Example

$$\begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 1 & 3 & 1 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \mathbf{1} \\ \mathbf{6} \\ \mathbf{7} \end{bmatrix} \quad \text{has the augmented matrix} \quad \begin{bmatrix} 1 & 3 & 0 & 2 & \mathbf{1} \\ 0 & 0 & 1 & 4 & \mathbf{6} \\ 1 & 3 & 1 & 6 & \mathbf{7} \end{bmatrix} = [A \;\; \boldsymbol{b}]$$

↓ row reduce

$$\begin{bmatrix} 1 & 3 & 0 & 2 & \mathbf{1} \\ 0 & 0 & 1 & 4 & \mathbf{6} \\ 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix} = [R \;\; \boldsymbol{d}].$$

Solve Ax=b

- Complete solution: $x = x_p + x_n$
- Particular solution x_p: $Ax_p = b$
  - Set free variables to be arbitrary values (e.g. zeros)
  - Solve for pivot variables (e.g. pivot variables from d)

$$\begin{bmatrix} 1 & 3 & 0 & 2 & \mathbf{1} \\ 0 & 0 & 1 & 4 & \mathbf{6} \\ 1 & 3 & 1 & 6 & \mathbf{7} \end{bmatrix} = \begin{bmatrix} A & \boldsymbol{b} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 0 & 2 & \mathbf{1} \\ 0 & 0 & 1 & 4 & \mathbf{6} \\ 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix} = \begin{bmatrix} R & \boldsymbol{d} \end{bmatrix}.$$

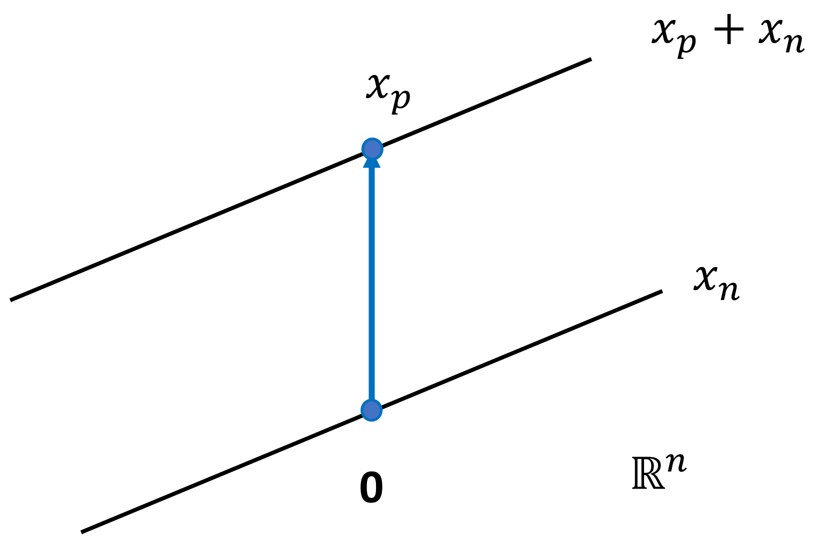
$$R\boldsymbol{x}_p = \begin{bmatrix} \mathbf{1} & 3 & \mathbf{0} & 2 \\ \mathbf{0} & 0 & \mathbf{1} & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{1} \\ 0 \\ \mathbf{6} \\ 0 \end{bmatrix}$$

- Free variables: x_2, x_4
- Pivot variables: x_1, x_3

Solve Ax=b

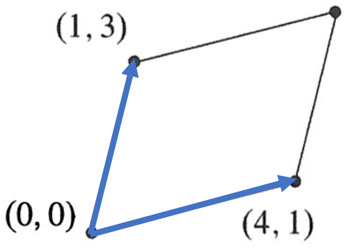
- Special solutions x_n: $Ax_n = \mathbf{0}$
  - $x_n = \mathrm{N(A)}$
  - $x_n = x_2 \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \forall x_2, x_4 \in \mathbb{R}$
- Complete solutions
  - Nullspace translated to pass point x_p

$$\boldsymbol{x} = \boldsymbol{x}_p + \boldsymbol{x}_n = \begin{bmatrix} 1 \\ 0 \\ 6 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 0 \\ -4 \\ 1 \end{bmatrix}$$



Determinant

- The determinant of a square matrix $A \in \mathbb{R}^{n\times n}$ is a function det: $\mathbb{R}^{n\times n} \to \mathbb{R}$, denoted as |A| or det(A).
- $|\det A|$ is the volume of the n-dimensional parallelotope formed by the row vectors of A
  - Example: $A = \begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix}$



$$\text{Area} = \begin{vmatrix} 4 & 1 \\ 1 & 3 \end{vmatrix} = 11$$

Determinant

- Let $A \in \mathbb{R}^{n \times n}$, properties of determinant:
  - $\det(I) = 1$
  - $\det(Q) = 1 \; or \; -1$
    - Q is an orthogonal matrix: $Q^T = Q^{-1}$
  - $\det(A) = 0 \iff rank(A) < n$
  - Let E be an elimination matrix, $\det(EA) = \det(E)\det(A)$
    - Multiply a row by a non-zero scalar t, det is multiplied by t
    - Subtracting a multiple of one row from another row leaves det(A) unchanged
    - det(A) changes sign when two rows are exchanged
  - $\det(A) = \det(A^T)$
  - If A is triangular then $\det(A) = \prod_{i=1}^{n} a_{ii}$
  - $\det(A) = \prod_{i=1}^{n} \lambda_i$

- Let $A, B \in \mathbb{R}^{n \times n}$, then $\det(AB) = \det(A)\det(B)$

Find det(A)

- Cofactor of element a_{ij}

$$C_{ij} = (-1)^{i+j} \det M_{ij}.$$

- Cofactor formula
  - det(A) is the dot product of ***any row i*** of A with its cofactors

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

- $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$
  - $\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$

Find det(A)

- Example

$$\det\begin{bmatrix} 2 & -3 & 1 \\ 2 & 0 & -1 \\ 1 & 4 & 5 \end{bmatrix} = 2\bullet\det\begin{bmatrix} 0 & -1 \\ 4 & 5 \end{bmatrix} - (-3)\bullet\det\begin{bmatrix} 2 & -1 \\ 1 & 5 \end{bmatrix} + 1\bullet\det\begin{bmatrix} 2 & 0 \\ 1 & 4 \end{bmatrix}$$

$$= 2\times 4 + 3\times 11 + 8 = 49$$

Eigenvalues and eigenvectors

- For a square matrix $A \in \mathbb{R}^{n \times n}$, $x \in \mathbb{R}^n (x \neq \mathbf{0})$ is called its eigenvector and $\lambda \in \mathbb{R}$ is its associated eigenvalue if
$$Ax = \lambda x$$
- An eigenvector x lies along the same line as Ax
- If $Ax = \lambda x$, then
  - $(A - \lambda I)x = \mathbf{0}$
  - $det(A - \lambda I) = 0$
- rank(A) is equal to the number of non-zero eigenvalues

Find eigenvalues

- Solve $det(A - \lambda I) = 0$ for λ
- $p(\lambda) = \ det(A - \lambda I)$ is the characteristic **polynomial of degree n** whose roots are eigenvalues
  - $det(A - \lambda I) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$

Find eigenvalues

Example

- $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

Subtract λ from the diagonal to find $A - \lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix}$

$$\det \begin{bmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix} = (1-\lambda)(4-\lambda) - (2)(2) = \lambda^2 - 5\lambda.$$

$\det(A - \lambda I) = \lambda^2 - 5\lambda = 0$ yields the eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 5$.

Find eigenvectors

- Solve $(A - \lambda I)x = \mathbf{0}$ for x
  - x is the $\mathrm{N}(A - \lambda I) - \{\mathbf{0}\}$
- Find the nullspace of $A - \lambda I$
  - $x = f(\lambda)$
- Plugging in each eigenvalue λ_i yields the corresponding eigenvector v_i

Find eigenvectors

Example

- $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

Subtract λ ***from the diagonal to find*** $A - \lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix}$

$$\begin{bmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix} \xrightarrow{R_2 \leftarrow R_2 - \frac{2}{1-\lambda} R_1} \begin{bmatrix} 1-\lambda & 2 \\ 0 & \frac{\lambda(5-\lambda)}{1-\lambda} \end{bmatrix} =$$

$$\xrightarrow{\lambda=0,5} \begin{bmatrix} 1-\lambda & 2 \\ 0 & 0 \end{bmatrix} \xrightarrow{R_1 \leftarrow \frac{1}{1-\lambda} R_1} \begin{bmatrix} 1 & \frac{2}{1-\lambda} \\ 0 & 0 \end{bmatrix}$$

Find eigenvectors

- $\begin{bmatrix} 1 & \frac{2}{1-\lambda} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
- $x = \begin{bmatrix} \frac{2}{\lambda-1} \\ 1 \end{bmatrix} x_2, \ \forall x_2 \in \mathbb{R}$ and $x_2 \neq 0$
- $\lambda_1 = 0$ gives the eigenvectors $x = c \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \ \forall c \in \mathbb{R}$ and c$\neq 0$
- $\lambda_2 = 5$ gives the eigenvectors $x = \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}, \forall c \in \mathbb{R}$ and c$\neq 0$

Inverse of a matrix

- For a square matrix $A \in \mathbb{R}^{n \times n}$
  - If A has an inverse (not singular), then $AA^{-1} = A^{-1}A = I$
  - Equivalent conditions of "A is invertible"
    - rank(A)=n
    - $det(A) \neq 0$
    - x=**0** is the only solution for $Ax = \mathbf{0}$
    - $Ax = b$ has a single solution $x = A^{-1}b$ for $\forall b \in \mathbb{R}$
    - All eigenvalues of A are nonzero

Find A^{-1} by elimination

- Gauss-Jordan Elimination
  - $[A \quad I] \xrightarrow{row\ reduction} [I \quad A^{-1}]$
  - $A^{-1}\,[A \quad I] = [I \quad A^{-1}]$
  - Can be applied to block matrix

Find A^{-1} by elimination

- Example

$$\begin{bmatrix} K & e_1 & e_2 & e_3 \end{bmatrix} = \left[\begin{array}{ccc|ccc} \mathbf{2} & \mathbf{-1} & \mathbf{0} & 1 & 0 & 0 \\ \mathbf{-1} & \mathbf{2} & \mathbf{-1} & 0 & 1 & 0 \\ \mathbf{0} & \mathbf{-1} & \mathbf{2} & 0 & 0 & 1 \end{array}\right] \quad \textbf{Start Gauss-Jordan on } K$$

$$\to \begin{bmatrix} 2 & -1 & 0 & 1 & 0 & 0 \\ \mathbf{0} & \mathbf{\tfrac{3}{2}} & \mathbf{-1} & \mathbf{\tfrac{1}{2}} & \mathbf{1} & \mathbf{0} \\ 0 & -1 & 2 & 0 & 0 & 1 \end{bmatrix} \quad (\tfrac{1}{2}\ \textit{row}\ 1 + \textit{row}\ 2)$$

$$\to \begin{bmatrix} 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & \tfrac{3}{2} & -1 & \tfrac{1}{2} & 1 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{\tfrac{4}{3}} & \mathbf{\tfrac{1}{3}} & \mathbf{\tfrac{2}{3}} & \mathbf{1} \end{bmatrix} \quad (\tfrac{2}{3}\ \textit{row}\ 2 + \textit{row}\ 3)$$

[U L]

Find A^{-1} by elimination

- Example

$$\begin{pmatrix}\text{Zero above}\\ \text{third pivot}\end{pmatrix} \to \left[\begin{array}{cccccc} 2 & -1 & 0 & 1 & 0 & 0 \\ \mathbf{0} & \frac{\mathbf{3}}{\mathbf{2}} & \mathbf{0} & \frac{\mathbf{3}}{\mathbf{4}} & \frac{\mathbf{3}}{\mathbf{2}} & \frac{\mathbf{3}}{\mathbf{4}} \\ 0 & 0 & \frac{4}{3} & \frac{1}{3} & \frac{2}{3} & 1 \end{array}\right] \qquad (\tfrac{\mathbf{3}}{\mathbf{4}}\ \boldsymbol{row}\ \mathbf{3} + \boldsymbol{row}\ \mathbf{2})$$

$$\begin{pmatrix}\text{Zero above}\\ \text{second pivot}\end{pmatrix} \to \left[\begin{array}{cccccc} \mathbf{2} & \mathbf{0} & \mathbf{0} & \frac{\mathbf{3}}{\mathbf{2}} & \mathbf{1} & \frac{\mathbf{1}}{\mathbf{2}} \\ 0 & \frac{3}{2} & 0 & \frac{3}{4} & \frac{3}{2} & \frac{3}{4} \\ 0 & 0 & \frac{4}{3} & \frac{1}{3} & \frac{2}{3} & 1 \end{array}\right] \qquad (\tfrac{\mathbf{2}}{\mathbf{3}}\ \boldsymbol{row}\ \mathbf{2} + \boldsymbol{row}\ \mathbf{1})$$

[D M]

Find A^{-1} by elimination

- Example

$$\begin{matrix}(\text{divide by } 2)\\(\text{divide by } \frac{3}{2})\\(\text{divide by } \frac{4}{3})\end{matrix}\quad\left[\begin{array}{ccc|ccc}1 & 0 & 0 & \frac{3}{4} & \frac{1}{2} & \frac{1}{4}\\ 0 & 1 & 0 & \frac{1}{2} & 1 & \frac{1}{2}\\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{2} & \frac{3}{4}\end{array}\right] = \begin{bmatrix} I & \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_3\end{bmatrix} = \begin{bmatrix} I & \boldsymbol{K}^{-1}\end{bmatrix}$$

Find A^{-1} by the adjoint of A

- $A^{-1} = \frac{1}{\det(A)} adj(A)$
- Find det(A)
- Find the adjoint of matrix A: $adj(A)$
  - $adj(A) = C^T$
  - $C = [C_{ij}]_{n \times n}$ is the cofactor matrix

Find A^{-1} by the adjoint of A

- Example
  - $A = \begin{bmatrix} 2 & 4 \\ 2 & 1 \end{bmatrix}$
  - $\det(A) = 2\times 1 - 2\times 4 = -6$
  - $C = \begin{bmatrix} 1 & -2 \\ -4 & 2 \end{bmatrix}$
  - $adj(A) = C^T = \begin{bmatrix} 1 & -4 \\ -2 & 2 \end{bmatrix}$
  - $A^{-1} = \frac{1}{\det(A)} adj(A) = \begin{bmatrix} -1/6 & 2/3 \\ 1/3 & -1/3 \end{bmatrix}$

Trace

- The trace of an $n \times n$ square matrix A is defined to be the sum of its diagonal elements
  - $tr(A) = \sum_{i=1}^{n} A_{ii}$
- Properties
  - For $A \in \mathbb{R}^{n \times n}$
    - $tr(A) = tr(A^T)$
    - $\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i$
    - $\mathrm{c} \in \mathbb{R}, tr(cA) = c\ tr(A)$
  - For $A, B \in \mathbb{R}^{n \times n}$
    - $tr(A + B) = tr(A) + tr(B)$
    - $tr(AB) = tr(BA)$

Matrix and vector derivatives: Gradient

Gradient: consider a function $f : \mathbb{R}^n \to \mathbb{R}$ that takes a vector $x \in \mathbb{R}^n$ and returns a real value. Then the gradient of f (w.r.t. x) is the vector of partial derivatives, defined as

$$\nabla_x f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix}.$$

- Examples
  - $f(x) = \mathbf{1}^T x = \sum_{i=0}^{n} x_i$
    - $\nabla f(x) = \mathbf{1}$
  - $f(x) = x^T x = \sum_{i=0}^{n} x_i^2$
    - $\nabla f(x) = 2x$

Matrix and vector derivatives : Gradient

Gradient: consider a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ that takes a vector $x \in \mathbb{R}^n$ and returns a real value. Then the gradient of f (w.r.t. x) is the vector of partial derivatives, defined as

$$\nabla_x f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix}.$$

- Examples
  - $f(x) = b^T x = \sum_{i=0}^{n} b_i x_i$
    - $\nabla f(x) = b$
  - $f(x) = x^T A x = \sum_{j=1}^{n} \sum_{i=1}^{n} x_j A_{ji} x_i$
    - $\nabla f(x) = (A + A^T)x$

Matrix and vector derivatives: Hessian

Hessian: consider a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ that takes a vector $x \in \mathbb{R}^n$ and returns a real value. Then the Hessian matrix of f (w.r.t. x) is the $n \times n$ matrix of partial derivatives, defined as

$$\nabla_x^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}.$$

- Examples
  - $f(x) = b^T x = \sum_{i=0}^{n} b_i x_i$
    - $\nabla^2 f(x) = [0]_{n \times n}$
  - $f(x) = x^T A x = \sum_{j=1}^{n} \sum_{i=1}^{n} x_j A_{ji} x_i$
    - $\nabla^2 f(x) = \mathrm{A} + \mathrm{A}^{\mathrm{T}}$