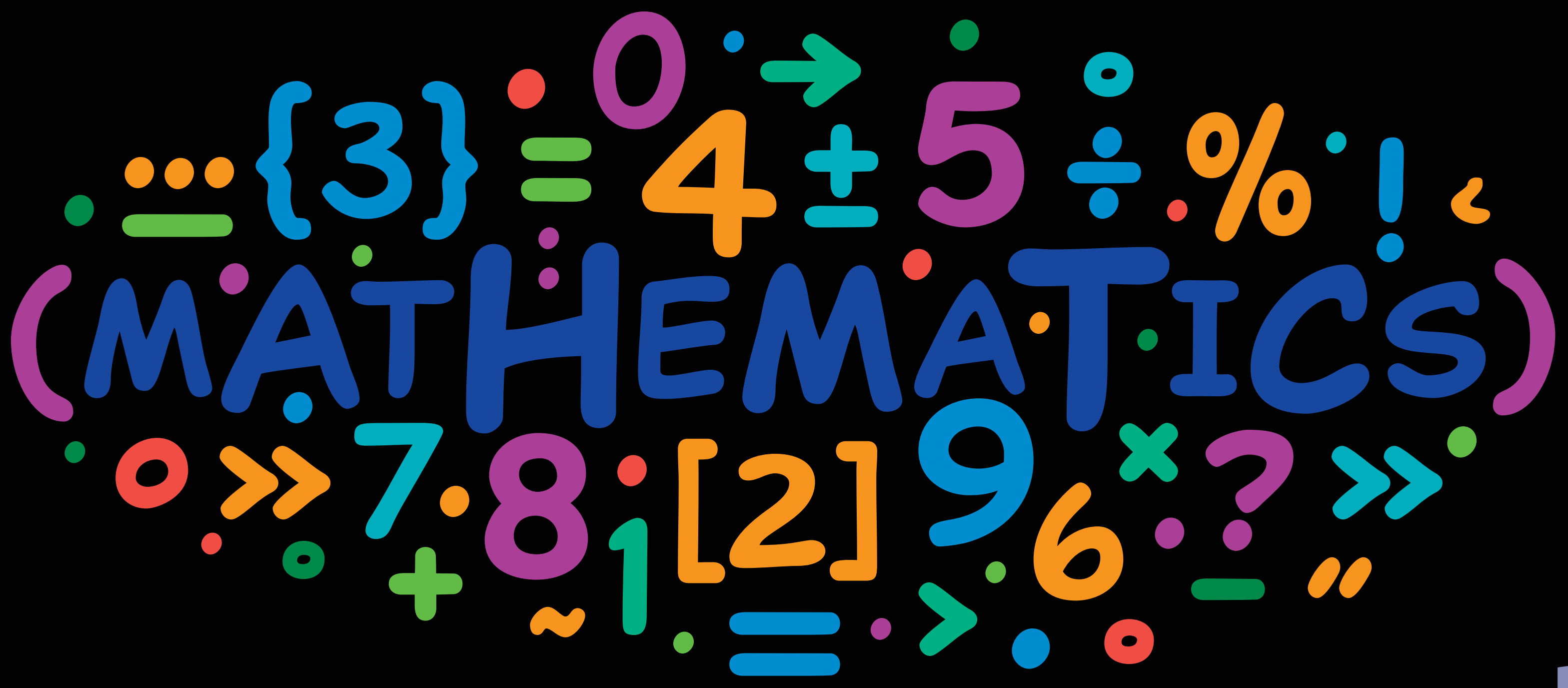
מתמטיקה ללמידת מכונה (22964)

מרכז הוראה: ד"ר שי מימון



פונקציות לינאריות של וקטורים

Linear and Affine Functions

- Linear functions – properties
- Linear functions and inner products
- Affine functions – properties
- A civil engineering example

Functions of Vectors

- The notation $f: \mathcal{R}^n \rightarrow R$ means that f is a function that maps real n-vectors to real numbers, i.e., it is a scalar-valued function of n-vectors
- If $\boldsymbol{x}$ is an n - vector, then $f(\boldsymbol{x})$, which is a scalar, denotes the value of the function f at $\boldsymbol{x}$

Linear Functions

A function $f: \mathcal{R}^n \rightarrow \mathcal{R}$ is **linear** if it satisfies the following two properties:

Homogeneity: For any n-vector $\boldsymbol{x}$ and any scalar α, $f(\alpha\boldsymbol{x}) = \alpha f(\boldsymbol{x})$

Additivity: For any n-vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, $f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x})+f(\boldsymbol{y})$

Linear Functions

These two properties (homogeneity and additivity) are equivalent to the **superposition** property:

$$f(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) = \alpha f(\boldsymbol{x}) + \beta f(\boldsymbol{y})$$

for any n-vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ and any scalars α and β

Linear Functions

If a function f is linear, superposition extends to linear combinations of any number of vectors, and not just linear combinations of two vectors:

$$f(\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_k \boldsymbol{x}_k) = \alpha_1 f(\boldsymbol{x}_1) + \cdots + \alpha_k f(\boldsymbol{x}_k),$$

for any n-vectors $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_k$, and any scalars $\alpha_1,\ldots,\alpha_k$

The Superposition Property

- For any n-vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ and any scalars α and β :

$$f(\alpha\boldsymbol{x}+\beta\boldsymbol{y})=\alpha f(\boldsymbol{x})+\beta f(\boldsymbol{y})$$

- **Homogeneity:** $\beta = 0 \rightarrow f(\alpha\boldsymbol{x}) = \alpha f(\boldsymbol{x})$

- **Additivity:** $\alpha = 1, \beta=1 \rightarrow f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$

The Inner Product Function

The inner product of the n-vector argument $\boldsymbol{x}$ with a fixed n-vector $\boldsymbol{a}$ is

$$f(\boldsymbol{x}) = \boldsymbol{a}^T \boldsymbol{x} = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$

The inner product with a fixed vector is a linear function

$$\begin{aligned} f(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) &= \boldsymbol{a}^T(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) \\ &= \boldsymbol{a}^T(\alpha \boldsymbol{x}) + \boldsymbol{a}^T(\beta \boldsymbol{y}) \\ &= \alpha(\boldsymbol{a}^T \boldsymbol{x}) + \beta(\boldsymbol{a}^T \boldsymbol{y}) \\ &= \alpha f(\boldsymbol{x}) + \beta f(\boldsymbol{y}) \end{aligned}$$

Inner Product Representation of a Linear Function

- **If a function is linear, then it can be expressed as the inner product of its argument with some fixed vector**
- If f is **linear,** then by multi-term superposition we have

$$f(\boldsymbol{x}) = f(x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n) = x_1\, f(\boldsymbol{e}_1) + \cdots + x_n\, f(\boldsymbol{e}_n) = \boldsymbol{a}^T\boldsymbol{x}$$

with $\boldsymbol{a} = \big(f(\boldsymbol{e}_1), f(\boldsymbol{e}_2), \ldots, f(\boldsymbol{e}_n)\big)$

- The representation of a linear function f as $f(\boldsymbol{x})=\boldsymbol{a}^T\boldsymbol{x}$ is **unique**

Affine functions

- A function $f: \mathcal{R}_n \rightarrow \mathcal{R}$ is **affine** if and only if it can be expressed as

  $f(\boldsymbol{x}) = \boldsymbol{a}^T\boldsymbol{x} + b$ for some n-vector $\boldsymbol{a}$ and scalar b
- Any affine scalar-valued function satisfies the following variation on the superposition property:

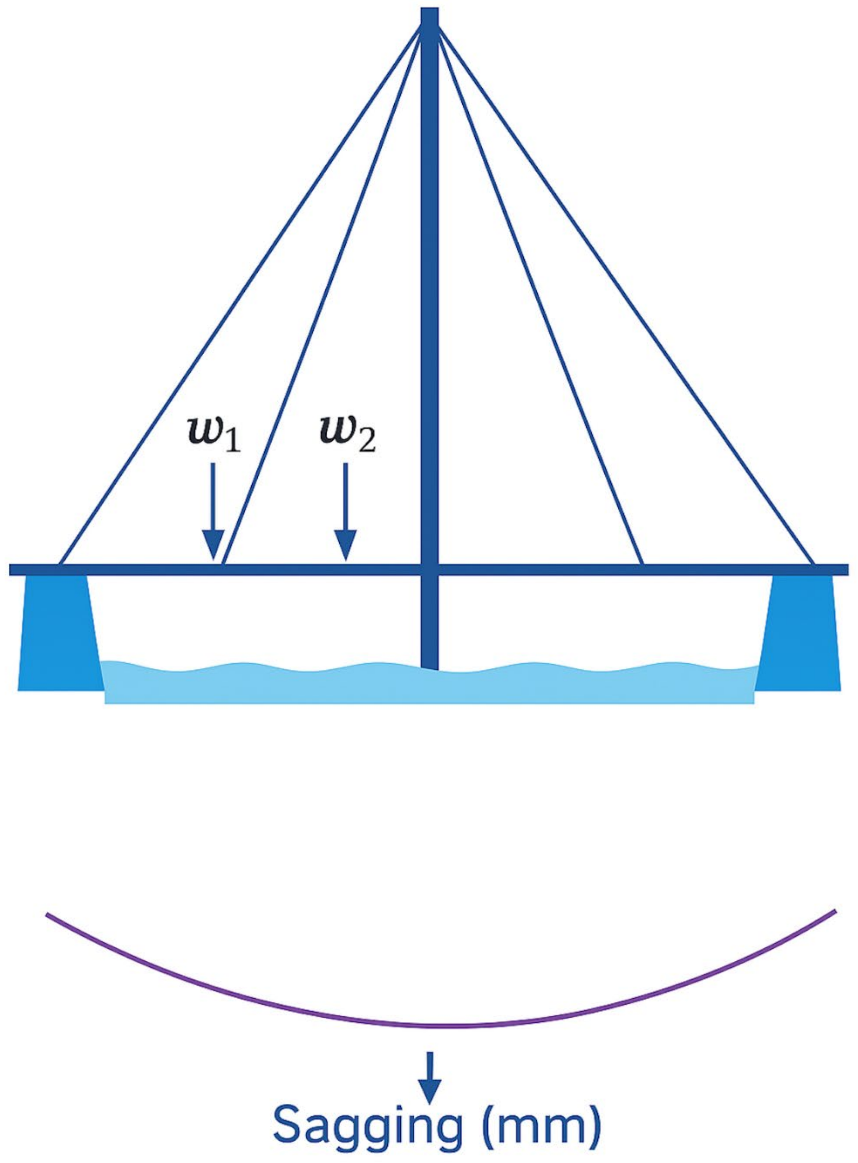
  $$f(\alpha\boldsymbol{x}+\beta\boldsymbol{y}) = \alpha f(\boldsymbol{x}) + \beta f(\boldsymbol{y}),$$

  for all n-vectors $\boldsymbol{x}$, $\boldsymbol{y}$, and all scalars α, β that satisfy $\alpha+\beta = 1$
- The converse is also true: Any scalar-valued function that satisfies the restricted superposition property is affine
- When f is **affine**, and $\boldsymbol{x}$ is any n-vector,

  $$f(\boldsymbol{x}) = f(\boldsymbol{0}) + x_1\,(f(\boldsymbol{e}_1) - f(\boldsymbol{0})) + \cdots + x_n\,(f(\boldsymbol{e}_n) - f(\boldsymbol{0})).$$

A Civil Engineering Example

- Consider a bridge and let $\boldsymbol{w}$ be an n-vector that gives the weight of the load on the bridge in n specific locations, in metric tons
- Let s denote the vertical distance that a specific point on the bridge sags (sinks), in millimeters, due to the load $\boldsymbol{w}$
- For weights the bridge is designed to handle, the sag is very well approximated as a linear function: $s = \boldsymbol{c}^T \boldsymbol{w}$, for some n-vector $\boldsymbol{c}$



A Civil Engineering Example

The vector $\boldsymbol{c}$ can be measured once the bridge is built

- We apply the load $\boldsymbol{w} = \boldsymbol{e}_1$ and measure the sag, which is c_1
- We repeat this experiment, moving the one ton load to positions 2,3,…,n, which gives us the coefficients $c_2,\ldots,c_n$
- At this point we have the vector $\boldsymbol{c}$, so we can now predict what the sag will be with any other loading

w_1	w_2	w_3	Measured sag	Predicted sag
1	0	0	0.12	—
0	1	0	0.31	—
0	0	1	0.26	—
0.5	1.1	0.3	0.481	0.479
1.5	0.8	1.2	0.736	0.740

Summary

- Linear functions:
  - ✓ Homogenity
  - ✓ Additivity
  - ✓ The superposition principle

- Affine functions: restricted superposition
- Bridge load and sag analysis