Modern Methods of Statistical Learning sf2935: Vector Spaces Timo Koski

TK

2017-11-01

KTH Matematik

Prerequisites on Vector Spaces

A set $\mathbf{X}$ with elements $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ ..., referred to as *vectors*, is called a *vector space*, if there are two operations called 'addition of two vectors' and 'multiplication of a vector by scalar', $\alpha \in R$. These operations satisfy

- $\mathbf{x} + \mathbf{y} \in \mathbf{X}$
- There is a neutral element $\mathbf{0}$ such that $\mathbf{x} + \mathbf{0} = \mathbf{x}$
- $\alpha \mathbf{x} \in \mathbf{X}$
- $1\mathbf{x} = \mathbf{x} \in \mathbf{X}$, $0\mathbf{x} = \mathbf{0} \in \mathbf{X}$

Prerequisites (A) on Vector Spaces

and in addition

- $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
- $\alpha \left(\mathbf{x} + \mathbf{y} \right) = \alpha\mathbf{x} + \alpha\mathbf{y}$
- $\left(\alpha + \beta \right) \mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$

We take

$$\mathbf{x} - \mathbf{y} = \mathbf{x} + (-1)\mathbf{y}$$

and then

$$\mathbf{x} - \mathbf{x} = \mathbf{0}$$

KTH Matematik

Prerequisites (A) on Vector Spaces: An Example

The standard example is the set R^n of real column vectors of fixed dimension n. Let ^{T} denote transpose

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$$

(a transposed row vector is a column vector), where
$x_i \in R, \quad i = 1, \ldots, n.$

$$\mathbf{0} = (0, 0, \ldots, 0)^T$$

We define

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)^T$$

$$\alpha \mathbf{x} = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n)^T$$

KTH Matematik

Prerequisites (A): Vector Spaces and Norms

A vector space $\mathbf{X}$ is called a *normed linear space*, if there is a real-valued function that maps each $\mathbf{x} \in \mathbf{X}$ to a number $\|\mathbf{x}\|$ with the following properties

- $\|\mathbf{x}\| \geq 0$ for all $\mathbf{x} \in \mathbf{X}$
- $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$.
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (triangle inequality)
- $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$ (homogeneity)

Linear Vector Spaces and Norms: Examples

Consider R^n. Then $\|\mathbf{x}\|_2$ and $\|\mathbf{x}\|_\infty$ are norms on R^n:

- $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$ (Euclidean norm)
- $\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$

Prerequisites (A): Norms and Distances

In a normed linear space $\mathbf{X}$ the real-valued function

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$

is called the *distance* between $\mathbf{x}$ and $\mathbf{y}$.

- $d(\mathbf{x}, \mathbf{y}) \geq 0$ for all $\mathbf{x} \in \mathbf{X}$
- $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$.
- $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$.
- $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{y}, \mathbf{z})$ (triangle inequality)
- $d(\alpha\mathbf{x}, \alpha\mathbf{y}) = |\alpha| d(\mathbf{x}, \mathbf{y})$ (homogeneity)

KTH Matematik

Lengths and Balls in Normed Spaces

Then the *length* of $\mathbf{x}$ is the distance from $\mathbf{x}$ to $\mathbf{0}$, i.e.,

$$d(\mathbf{x}, \mathbf{0}) = \|\mathbf{x} - \mathbf{0}\| = \|\mathbf{x}\|$$

The **open ball** $\mathcal{B}_\tau(\mathbf{x}) \subset X$ of radius τ around $\mathbf{x} \in X$ is

$$\mathcal{B}_\tau(\mathbf{x}) \stackrel{\text{def}}{=} \{\mathbf{y} \in X \mid \|\mathbf{y} - \mathbf{x}\| < \tau\}$$

Prerequisites (A): Inner Product Spaces

A vector space $\mathbf{X}$ is called an *inner product space*, if there is a function, called inner product[1], that maps each pair $\mathbf{x}, \mathbf{y}$ of vectors in $\mathbf{X}$ to a number $\ll \mathbf{x}, \mathbf{y} \gg$ with the following properties

- $\ll \mathbf{x}, \mathbf{y} \gg = \ll \mathbf{y}, \mathbf{x} \gg$
- $\ll \mathbf{x}, \mathbf{x} \gg \geq 0$ with $\ll \mathbf{x}, \mathbf{x} \gg = 0$, if and only if $\mathbf{x} = \mathbf{0}$.
- $\ll \mathbf{x} + \mathbf{y}, \mathbf{z} \gg = \ll \mathbf{x}, \mathbf{z} \gg + \ll \mathbf{y}, \mathbf{z} \gg$
- $\ll \mathbf{x}, \mathbf{y} + \mathbf{z} \gg = \ll \mathbf{x}, \mathbf{y} \gg + \ll \mathbf{x}, \mathbf{z} \gg$
- $\ll \mathbf{x}, \alpha\mathbf{y} \gg = \alpha \ll \mathbf{x}, \mathbf{y} \gg$

[1] Much of learning theory literature talks about dot products

Inner Product Spaces: Examples

1. Take $\mathbf{X} = R^n$, and

$$\ll \mathbf{x}, \mathbf{y} \gg = \sum_{i=1}^{n} x_i y_i$$

2. Take $\mathbf{X} = R^n$, and

$$\ll \mathbf{x}, \mathbf{y} \gg_A = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i y_j = \ll \mathbf{x}, A\mathbf{y} \gg$$

where $A = (a_{ij})_{i,j=1}^{n.n}$ is a symmetric and non-negative definite matrix[2]

[2] $\ll \mathbf{x}, A\mathbf{y} \gg = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \geq 0$ for all $\mathbf{x}$.

Inner Product Spaces are Normed Spaces

An inner product space $\mathbf{X}$ is automatically a normed linear space, since we can put

$$\|\mathbf{x}\| = \sqrt{\ll \mathbf{x}, \mathbf{x} \gg}$$

and, since $\ll \mathbf{x}, \mathbf{x} \gg \geq 0$,

$$\|\mathbf{x}\|^2 = \ll \mathbf{x}, \mathbf{x} \gg$$

Example: $\mathbf{X} = R^n$

$$\|\mathbf{x}\| = \sqrt{\ll \mathbf{x}, \mathbf{x} \gg} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Examples: Mahalanobis distance

When we take $A = \Sigma^{-1}$ in $\ll \mathbf{x}, \mathbf{y} \gg_A$ and set

$$\|\mathbf{x} - \mathbf{y}\|^2_{\Sigma^{-1}} = \ll \mathbf{x} - \mathbf{y}, \Sigma^{-1} (\mathbf{x} - \mathbf{y}) \gg$$

we obtain a useful distance in pattern recognition known as the (squared) *Mahalanobis distance*. Regions of constant Mahalanobis distance to a fixed vector $\mathbf{y}$ are ellipsoids.

Properties of Inner Products

- Cauchy-Schwarz inequality

$$|\ll \mathbf{x}, \mathbf{y} \gg| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$$

- $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2 \ll \mathbf{x}, \mathbf{y} \gg$
- $\|\mathbf{x}-\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2 \ll \mathbf{x}, \mathbf{y} \gg$

Properties of Inner Product Spaces

The angle θ between $\mathbf{x}$ and $\mathbf{y}$ in an inner product space is given by

- $\cos(\theta) = \frac{\ll \mathbf{x}, \mathbf{y} \gg}{\|\mathbf{x}\| \|\mathbf{y}\|}$

If $\ll \mathbf{x}, \mathbf{y} \gg = 0$, we say that $\mathbf{x}$ and $\mathbf{y}$ are *orthogonal*, since then $\cos(\theta) = 0$, and then $\theta = \pi/2$ (within period). Also, we have then the *Pythagorean relations*

- $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$
- $\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$

KTH Matematik

Orthonormal Vectors in Inner Product Spaces

Let $\phi_1, \ldots, \phi_n$ be a sequence of **orthonormal** vectors of an inner product space X. Orthonormality means that

1. $\ll \phi_i, \phi_j \gg = 0$, for $i \neq j$.
2. $\|\phi_i\| = 1$ for all j.

We note an example.

1. $\mathbf{X} = R^n$, and $\ll \mathbf{x}, \mathbf{y} \gg = \sum_{i=1}^{n} x_i y_i$. Then

$$\phi_i = \left(0, 0, \ldots, \underbrace{1}_{i\text{th position}}, \ldots, 0 \right), \quad i = 1, \ldots, n$$

are orthonormal vectors in $\mathbf{X} = R^n$.

Basis and Dimension

We say that the orthonomal vectors $\phi_1, \ldots, \phi_n$ form an orthonormal *basis* for X, if every $\mathbf{x}$ can be written

$$\mathbf{x} = \ll \mathbf{x}, \phi_1 \gg \phi_1 + \ldots + \ll \mathbf{x}, \phi_n \gg \phi_n$$

and that this is unique, also in the sense $\phi_1, \ldots, \phi_n$ cannot be appended or reduced by a single orthonormal vector. Then n is the *dimension* of the inner product space X.

Inner Product Spaces

An inner product space supports thus geometric notions. Inner product spaces of finite dimension are often called Euclidean spaces. A Euclidean space X equipped with $\ll \mathbf{x}, \mathbf{y} \gg_X$ is *isomorphic* to R^n with $\ll \mathbf{x}, \mathbf{y} \gg = \sum_{i=1}^n x_i y_i$ in the sense that there is an invertible map $\psi : X \mapsto R^n$ such that

$$\ll \mathbf{x}, \mathbf{y} \gg_X = \ll \psi(\mathbf{x}), \psi(\mathbf{y}) \gg$$

KTH Matematik