Data Science for Actuaries (ACT6100)

Arthur Charpentier

Rappels # 1 (Vectors, Norms & Inner Product)

automne 2020

https://github.com/freakonometrics/ACT6100/

Norm

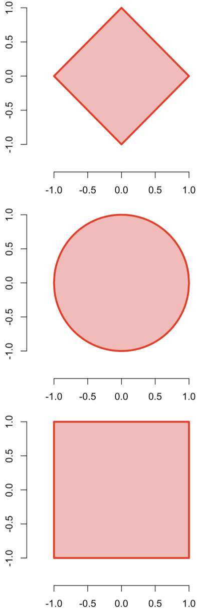
A norm $\|\cdot\|$, in $\mathbb{R}^n$, satisfies

- homogeneity, $\|a\vec{\boldsymbol{u}}\| = |a| \cdot \|\vec{\boldsymbol{u}}\|$
- triangle inequality, $\|\vec{\boldsymbol{u}} + \vec{\boldsymbol{v}}\| \leq \|\vec{\boldsymbol{u}}\| + \|\vec{\boldsymbol{v}}\|$
- positivity, $\|\vec{\boldsymbol{u}}\| \geq 0$
- definiteness, $\|\vec{\boldsymbol{u}}\| = 0 \iff \vec{\boldsymbol{u}} = \vec{\boldsymbol{0}}$

ℓ_1 norm: $\|\boldsymbol{x}\|_{\ell_1} = |x_1| + \cdots + |x_n|$,
see taxicab geometry
ℓ_p norm: with $p \geq 1$,

$$\|\boldsymbol{x}\|_{\ell_p} = (|x_1|^p + \cdots + |x_n|^p)^{1/p}$$

e.g. $\|\boldsymbol{x}\|_{\ell_\infty} = \max\{x_i\}$
Unit balls ($\|\boldsymbol{x}\| \leq 1$) are convex sets



Hilbert Space and Inner Products

An inner product $\langle\cdot,\cdot\rangle$, in $\mathbb{R}^n$, satisfies

- symmetry, $\langle\vec{\boldsymbol{u}},\vec{\boldsymbol{v}}\rangle=\langle\vec{\boldsymbol{v}},\vec{\boldsymbol{u}}\rangle$
- linearity, $\langle a\vec{\boldsymbol{u}}+b\vec{\boldsymbol{v}},\vec{\boldsymbol{w}}\rangle=a\langle\vec{\boldsymbol{u}},\vec{\boldsymbol{w}}\rangle+b\langle\vec{\boldsymbol{v}},\vec{\boldsymbol{w}}\rangle$
- positivity, $\langle\vec{\boldsymbol{u}},\vec{\boldsymbol{u}}\rangle\geq 0$
- definiteness, $\langle\vec{\boldsymbol{u}},\vec{\boldsymbol{u}}\rangle=0 \iff \vec{\boldsymbol{u}}=\vec{\boldsymbol{0}}$

Example: On the set of $\mathbb{R}^n$ vectors, $\langle\boldsymbol{x},\boldsymbol{y}\rangle=\boldsymbol{x}^\top\boldsymbol{y}$

Furthermore

- $\|\boldsymbol{x}\|=\sqrt{\langle\boldsymbol{x},\boldsymbol{x}\rangle}$ defines a norm
- $d(\boldsymbol{x},\boldsymbol{y})=\|\boldsymbol{x}-\boldsymbol{y}\|$ defines a distance

Example: On the set of $m\times n$ matrices, $\langle\boldsymbol{A},\boldsymbol{B}\rangle=\text{trace}\big(\boldsymbol{A}\boldsymbol{B}^\top\big)$

Example: On the set of ranfom variables, $\langle X,Y\rangle=\mathbb{E}(XY)$

Cauchy-Schwarz Inequality

$$|\langle \boldsymbol{x}, \boldsymbol{y}\rangle| \leq \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\| \text{ with equality only when } \boldsymbol{x} = \lambda\boldsymbol{y}$$

Application: $x_i \leftarrow x_i - \overline{x}$ and $y_i \leftarrow y_i - \overline{y}$,

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \boldsymbol{x}^\top \boldsymbol{y} = \sum_{i=1}^n x_i y_i$$

$$\|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\| = \sqrt{\boldsymbol{x}^\top \boldsymbol{x}} \cdot \sqrt{\boldsymbol{y}^\top \boldsymbol{y}} = \sqrt{\sum_{i=1}^n x_i^2} \cdot \sqrt{\sum_{i=1}^n y_i}$$

$$\text{corr}(\boldsymbol{x}, \boldsymbol{y}) = \frac{\displaystyle\sum_{i=1}^n x_i y_i}{\displaystyle\sum_{i=1}^n x_i^2 \cdot \sum_{i=1}^n y_i^2} = \frac{\langle \boldsymbol{x}, \boldsymbol{y}\rangle}{\|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|} \in [-1, +1]$$

and $\text{corr}(\boldsymbol{x}, \boldsymbol{y}) = \pm 1$ only when $\boldsymbol{x} = \lambda \boldsymbol{y}$.

Mahalanobis distance

A $n \times n$ symmetric matrix $\boldsymbol{M}$ is positive definite if $\boldsymbol{x}^\top \boldsymbol{M}\boldsymbol{x} > 0$ for any $\boldsymbol{x} \in \mathbb{R}^n$.

Proposition: If $\boldsymbol{M}$ is a positive definite (symmetric) matrix, then $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^\top \boldsymbol{M}\boldsymbol{y}$ defines an inner product.

(furthermore, conversely, if $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^\top \boldsymbol{M}\boldsymbol{y}$ defines an inner product, then $\boldsymbol{M}$ is definite positive)

- $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_M = \boldsymbol{x}^\top \boldsymbol{M}\boldsymbol{y}$ defines an inner product,
- $\|\boldsymbol{x}\|_M = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle_M}$ defines a norm
- $d_M(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|_M$ defines a distance

Given $\boldsymbol{\Sigma}$ some $n \times n$ definite positive matrix, define

$$d(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{(\boldsymbol{x} - \boldsymbol{y})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{y})}$$

Given $\boldsymbol{\mu} \in \mathbb{R}^n$, define the Mahalanobis "norm"

$$\|\boldsymbol{x}\| = d(\boldsymbol{x}, \boldsymbol{\mu}) = \sqrt{(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})}$$

If $\boldsymbol{\Sigma}$ is diagonal, it is also called standardized Euclidean distance.

See on the generalised distance in statistics, 1936.