Mathematics and Statistics (BEMACS) - A. Y. 2019/20

Lecture 3

BEMACS

Decision Sciences (Bocconi University)

11 September 2019

Cauchy-Schwarz inequality

Property 3.1. Let $\mathbf{x}^1$, $\mathbf{x}^2 \in \mathbb{R}^n$ be vectors. The inequality

$$\left|\left(\mathbf{x}^1, \mathbf{x}^2\right)\right| \leq \left\|\mathbf{x}^1\right\| \cdot \left\|\mathbf{x}^2\right\|$$

is called the Cauchy-Schwarz inequality.
Proof. If $\mathbf{x}^1 = \mathbf{0}$ or $\mathbf{x}^2 = \mathbf{0}$ then both sides of the inequality are 0. Otherwise, we introduce a real variable t and consider the vector $\mathbf{z} = t\mathbf{x}^1 + \mathbf{x}^2$. We have

$$\begin{aligned}
0 \quad \leq & \quad \|\mathbf{z}\|^2 = \left(\sqrt{(\mathbf{z}, \mathbf{z})}\right)^2 = (\mathbf{z}, \mathbf{z}) = \left(t\mathbf{x}^1 + \mathbf{x}^2, t\mathbf{x}^1 + \mathbf{x}^2\right) \\
= & \quad \left(t\mathbf{x}^1, t\mathbf{x}^1\right) + (t\mathbf{x}^1, \mathbf{x}^2) + (\mathbf{x}^2, t\mathbf{x}^1) + (\mathbf{x}^2, \mathbf{x}^2) \\
= & \quad t^2\left(\mathbf{x}^1, \mathbf{x}^1\right) + 2t\left(\mathbf{x}^1, \mathbf{x}^2\right) + (\mathbf{x}^2, \mathbf{x}^2) \\
= & \quad \left\|\mathbf{x}^1\right\|^2 t^2 + 2\left(\mathbf{x}^1, \mathbf{x}^2\right) t + \left\|\mathbf{x}^2\right\|^2
\end{aligned}$$

Cauchy-Schwarz inequality

Let $a = \left\|\mathbf{x}^1\right\|^2$, $b = \left(\mathbf{x}^1, \mathbf{x}^2\right)$ and $c = \left\|\mathbf{x}^2\right\|^2$. From $at^2 + 2bt + c \geq 0$, for any $t \in \mathbb{R}$. As a consequence, the discriminant $\Delta = 4b^2 - 4ac$ must be non-positive. Thus $b^2 \leq ac$. By substitution, we get

$$0 \leq \left|\left(\mathbf{x}^1, \mathbf{x}^2\right)\right|^2 \leq \left\|\mathbf{x}^1\right\|^2 \cdot \left\|\mathbf{x}^2\right\|^2 \quad \Rightarrow \quad 0 \leq \left|\left(\mathbf{x}^1, \mathbf{x}^2\right)\right| \leq \left\|\mathbf{x}^1\right\| \cdot \left\|\mathbf{x}^2\right\|$$

Suppose that $\mathbf{x}^2 = -t\mathbf{x}^1$, with $t \in \mathbb{R}$, that is $\mathbf{x}^1$ and $\mathbf{x}^1$ are collinear vectors on $\mathbb{R}^n$. The left side of the Cauchy-Schwarz inequality is

$$\left|\left(\mathbf{x}^1, \mathbf{x}^2\right)\right| = \left|\left(\mathbf{x}^1, -t\mathbf{x}^1\right)\right| = |-t| \left|\left(\mathbf{x}^1, \mathbf{x}^1\right)\right| = t \left\|\mathbf{x}^1\right\|^2$$

while the right one is

$$\left\|\mathbf{x}^1\right\| \cdot \left\|\mathbf{x}^2\right\| = \left\|\mathbf{x}^1\right\| \cdot \left\|-t\mathbf{x}^1\right\| = |-t| \left\|\mathbf{x}^1\right\| \left\|\mathbf{x}^1\right\| = t \left\|\mathbf{x}^1\right\|^2$$

and the Cauchy-Schwarz inequality holds with the equality sign.

Example 3.2. Consider on $\mathbb{R}^3$ the following vectors

$$\mathbf{x}^1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{x}^2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

The inner product of $\mathbf{x}^1$ and $\mathbf{x}^2$, the norm of $\mathbf{x}^1$ and the norm of $\mathbf{x}^2$ are

$$\left(\mathbf{x}^1, \mathbf{x}^2\right) = 1 \qquad \left\|\mathbf{x}^1\right\| = \sqrt{2} \qquad \left\|\mathbf{x}^2\right\| = \sqrt{2}$$

We have

$$\left|\left(\mathbf{x}^1, \mathbf{x}^2\right)\right| = 1 \leq 2 = \sqrt{2} \cdot \sqrt{2} = \left\|\mathbf{x}^1\right\| \cdot \left\|\mathbf{x}^2\right\|$$

Remark 3.3. A norm on $\mathbb{R}^n$ is any function from $\mathbb{R}^n$ to $\mathbb{R}$ that is positive definite, absolutely homogeneous and subadditive. The following functions from $\mathbb{R}^2$ to $\mathbb{R}$ defined by

$$\|\mathbf{x}\|_1 = |x_1| + |x_2| \qquad \|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2} \qquad \|\mathbf{x}\|_\infty = \max\{|x_1|, |x_2|\}$$

are norms on $\mathbb{R}^2$.

Norms on 2-dimension spaces

The 2-norm is the natural generalization of the Euclidean length of a 2-vector and is also called the Euclidean norm. The ∞-norm is sometimes called the maximum norm (max-norm) or the Chebyshev norm. All the three functions satisfy the three properties required to be a norm (i.e. positive definite, absolutely homogeneous and subadditive). Let's consider $\|\mathbf{x}\|_1 = |x_1| + |x_2|$. We have:

- For any $\mathbf{x} \in \mathbb{R}^2$, $\|\mathbf{x}\|_1 = |x_1| + |x_2| \geq 0$ and $\|\mathbf{x}\|_1 = 0$ if and only if $\mathbf{x} = \mathbf{0}$.
- For any $\mathbf{x} \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$

$$\|\alpha \mathbf{x}\|_1 = |\alpha x_1| + |\alpha x_2| = \alpha |x_1| + \alpha |x_2| = \alpha (|x_1| + |x_2|) = \alpha \|\mathbf{x}\|_1$$

- For any $\mathbf{x}^1, \mathbf{x}^2 \in \mathbb{R}^2$

$$\begin{aligned} \left\|\mathbf{x}^1 + \mathbf{x}^2\right\|_1 &= \left|x_1^1 + x_1^2\right| + \left|x_2^1 + x_2^2\right| \leq \left|x_1^1\right| + \left|x_1^2\right| + \left|x_2^1\right| + \left|x_2^2\right| \\ &= \left(\left|x_1^1\right| + \left|x_2^1\right|\right) + \left(\left|x_1^2\right| + \left|x_2^2\right|\right) = \left\|\mathbf{x}^1\right\|_1 + \left\|\mathbf{x}^2\right\|_1 \end{aligned}$$

Distance between two points and properties

Definition 3.4. Let $\mathbf{x}^1$, $\mathbf{x}^2 \in \mathbb{R}^n$ be vectors. The non-negative real number, indicated with $d\left(\mathbf{x}^1, \mathbf{x}^2\right)$, defined by

$$d\left(\mathbf{x}^1, \mathbf{x}^2\right) = \left\|\mathbf{x}^1 - \mathbf{x}^2\right\| = \sqrt{\sum_{s=1}^n \left(x_s^1 - x_s^2\right)^2}$$

is called the Euclidean distance between $\mathbf{x}^1$ and $\mathbf{x}^2$.

Property 3.5. Let $\mathbf{x}^1$, $\mathbf{x}^2$, $\mathbf{x}^3 \in \mathbb{R}^n$ be vectors. We have:

(a) $d(\mathbf{x}^1, \mathbf{x}^2) \geq 0$, $d(\mathbf{x}^1, \mathbf{x}^2) = 0$ if and only if $\mathbf{x}^1 = \mathbf{x}^2$ (non-negativity);

(b) $d(\mathbf{x}^1, \mathbf{x}^2) = d(\mathbf{x}^2, \mathbf{x}^1)$ (symmetry);

(c) $d(\mathbf{x}^1, \mathbf{x}^2) \leq d(\mathbf{x}^1, \mathbf{x}^3) + d(\mathbf{x}^3, \mathbf{x}^2)$ (triangle inequality), the equality sign holds if and only if $\mathbf{x}^3$ is a convex linear combination between $\mathbf{x}^1$ and $\mathbf{x}^2$.

Proof. (b) We have

$$d(\mathbf{x}^1, \mathbf{x}^2) = \left\|\mathbf{x}^1 - \mathbf{x}^2\right\| = \left\|-\left(\mathbf{x}^2 - \mathbf{x}^1\right)\right\| = \left\|\mathbf{x}^2 - \mathbf{x}^1\right\| = d(\mathbf{x}^2, \mathbf{x}^1)$$

Examples

Example 3.6. Consider on $\mathbb{R}^3$ the following vectors

$$\mathbf{x}^1 = \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} \quad ; \quad \mathbf{x}^2 = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} \quad ; \quad \mathbf{x}^3 = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}$$

We have the following distances between vectors $\mathbf{x}^1$, $\mathbf{x}^2$ and $\mathbf{x}^3$

$$d(\mathbf{x}^1, \mathbf{x}^2) = \left\| \mathbf{x}^1 - \mathbf{x}^2 \right\| = \sqrt{(-2+2)^2 + (1-1)^2 + (4-2)^2} = 2$$

and $d(\mathbf{x}^1, \mathbf{x}^3) = 1$, $d(\mathbf{x}^3, \mathbf{x}^2) = 1$. We observe that

$$d(\mathbf{x}^1, \mathbf{x}^2) = 2 \leq 1 + 1 = d(\mathbf{x}^1, \mathbf{x}^3) + d(\mathbf{x}^3, \mathbf{x}^2)$$

in fact $\mathbf{x}^3$ is a convex linear combination of $\mathbf{x}^1$, $\mathbf{x}^2$ (with both the coefficients $1/2$). But

$$d(\mathbf{x}^1, \mathbf{x}^3) = 1 \leq 2 + 1 = d(\mathbf{x}^1, \mathbf{x}^2) + d(\mathbf{x}^2, \mathbf{x}^3)$$

in fact $\mathbf{x}^2$ is not a convex linear combination of $\mathbf{x}^1$ and $\mathbf{x}^3$.

Neighbourhood of a point on n-dimension real spaces

Definition 3.7. Let $\mathbf{x}^0 \in \mathbb{R}^n$ be a vector. The set of points, indicated by $B_\varepsilon(\mathbf{x}^0)$, defined by

$$B_\varepsilon(\mathbf{x}^0) = \left\{\mathbf{x} \in \mathbb{R}^n : \mathrm{d}\left(\mathbf{x}, \mathbf{x}^0\right) < \varepsilon\right\}$$

is called the neighbourhood of $\mathbf{x}^0$ with radius $\varepsilon > 0$.

Remark 3.8. (a) From the definition of distance between two points, the previous definition is equivalent to

$$B_\varepsilon(\mathbf{x}^0) = \left\{\mathbf{x} \in \mathbb{R}^n : \left\|\mathbf{x} - \mathbf{x}^0\right\| < \varepsilon\right\}$$

(b) If $n = 1$, we have $B_\varepsilon(x^0) = \left(x^0 - \varepsilon, x^0 + \varepsilon\right)$ (i.e. an open interval); (c) If $n = 2$, we have

$$B_\varepsilon(\mathbf{x}^0) = \left\{\mathbf{x} \in \mathbb{R}^n : \left(x_1 - x_1^0\right)^2 + \left(x_2 - x_2^0\right)^2 < \varepsilon\right\}$$

which is equivalent to all the points inside the circle with center $\left(x_1^0, x_2^0\right)$ and radius $\varepsilon > 0$ (without the border).

Upper bound and lower bound of a set

Definition 3.9. Let $A \subseteq \mathbb{R}$ be a non-empty set. A number $h \in \mathbb{R}$, if it exists, such that

$$\forall x \in A, \quad x \leq h$$

is said to be an upper bound of A.

Definition 3.10. Let $A \subseteq \mathbb{R}$ be a non-empty set. A number $k \in \mathbb{R}$, if it exists, such that

$$\forall x \in A, \quad k \leq x$$

is said to be a lower bound of A.

From now on we will indicate with A^* the set of all the upper bounds of A and A_* the set of all the lower bounds of A.

Remark 3.11. If h (k) is an upper (lower) bound of A then every $h' < h$ ($k' < k$) will be an upper (lower) bound of A. Upper and lower bounds of a set A, if they exist, are not unique (i.e. there are infinite).

Example 3.12. Let $A = [0, 20)$ be a subset of $\mathbb{R}$. All the numbers greater than or equal to 20 are greater than or equal to all the elements of A, so A has at least an upper bound (e.g. $h = 20$). Any value $h' > 20$ is also an upper bound of A.

Examples

Example 3.13. Let $A = [2, +\infty)$ be a subset of $\mathbb{R}$. There is no number h such that all the elements of A are less than or equal to h, therefore A does not have an upper bound and we have $A^* = \emptyset$. All the numbers less than or equal to 2 are less than or equal to all the elements of A, so A has a lower bound (e.g. $k = 2$). We have $A_* = (-\infty, 2]$. Any value of A_* is a lower bound of A (i.e. infinite).

Example 3.14. Let $A = (-\infty, 10)$ and $B = (0, 10)$ be subsets of $\mathbb{R}$. We have

$$A_* = \emptyset,\ A^* = [10, +\infty) \qquad \text{and} \qquad B_* = (-\infty, 0]\,,\ B^* = [10, +\infty)$$

Example 3.15. Let $A = \{x \in \mathbb{R} : |x| \leq 4\}$ be a subset of $\mathbb{R}$. We have

$$|x| \leq 4 \quad \Rightarrow \quad -4 \leq x \leq 4 \quad \Rightarrow \quad A = [-4, 4]$$

therefore $A_* = (-\infty, -4]$, $A^* = [4, +\infty)$. The complement of A is

$$A = \{x \in \mathbb{R} : |x| > 4\} \qquad \text{or} \qquad A = (-\infty, -4) \cup (4, +\infty)$$

therefore $A_* = A^* = \emptyset$.

Maximum and minimum of a set

Definition 3.16. Let $A \subseteq \mathbb{R}$ be a non–empty set. A is said to be:

1. bounded from above if it has an upper bound ($A^* \neq \emptyset$);
2. bounded from below if it has a lower bound ($A_* \neq \emptyset$);
3. bounded if it is bounded from above and below.

Example 3.17. Let $A = (0, +\infty)$, $B = (-\infty, 2]$ and $C = [-2, 2]$. We have

$$A_* = (-\infty, 0] \qquad B^* = [2, +\infty) \qquad C_* = (-\infty, -2] \qquad C^* = [2, +\infty)$$

therefore A is bounded from below, B is bounded from above and C is bounded.

Definition 3.18. Let $A \subseteq \mathbb{R}$ be a non–empty set. A number $x^* \in A$, if it exists, is called a:

1. maximum of A if $\forall x \in A$, $x \leq x^*$;
2. minimum of A if $\forall x \in A$, $x^* \leq x$.

Maximum an minimum of any set must belong to A (if they exist).

Maximum and minimum of a set

Theorem 3.19. Let $A \subseteq \mathbb{R}$ be a non-empty set. A has, if they exist, at most a maximum and minimum.

Example 3.20. Let $B = (-\infty, 2]$ be a subset of $\mathbb{R}$. We have

$$B_* = \emptyset \qquad B^* = [2, +\infty) \qquad \nexists \min B \qquad \max B = 2$$

and B is bounded from above. If we consider $B^c = (2, +\infty)$, we have

$$B^c_* = (-\infty, 2] \qquad B^{c*} = \emptyset \qquad \nexists \min B^c \qquad \nexists \max B^c$$

Remark 3.21. The maximum and minimum of A, if they exist, are

$$\min A = \max A_* \qquad \text{and} \qquad \max A = \min A^*$$

As previously seen, the maximum and minimum of a set cannot exist. So we can introduce the definition of two new mathematical objects: supremum and infimum of a set.

Supremum and infimum of a set

Definition 3.22. Let $A \subseteq \mathbb{R}$ be a non-empty set.

1. The least upper bound of A, if it exists, is called the supremum of A;
2. The greatest lower bound of A, if it exists, is called the infimum of A.

Example 3.23. Let $B = (-\infty, 2]$ be a subset of $\mathbb{R}$. We have

$$\sup B = 2 \qquad \text{and} \qquad \nexists \ \inf B$$

Remark 3.24. As you know we have to distinguish two different working environments: $\mathbb{R}$ (the real line) and $\overline{\mathbb{R}}$ (the extended real line). Therefore:

- in $\mathbb{R}$ the supremum and infimum of A cannot exist, therefore we write $\nexists \sup A$ (i.e. $\sup A$ does not exist) and $\nexists \inf A$ (i.e. $\inf A$ does not exist);
- in $\overline{\mathbb{R}}$ the supremum and infimum are $\sup A = +\infty$ and $\inf A = -\infty$.

So be careful on which environment you are working.

Example 3.25. Let $A = [0, +\infty)$ be a subset of $\mathbb{R}$. If we are working on $\mathbb{R}$ we have

$$\nexists \max A \qquad A^* = \emptyset \qquad \nexists \sup A$$

If we are working on $\overline{\mathbb{R}}$ we have

$$\nexists \max A \qquad A^* = \emptyset \qquad \sup A = +\infty$$

Example 3.26. Let $A = (-\infty, 2)$ be a subset of $\mathbb{R}$. If we are working on $\mathbb{R}$ we have

$$\nexists \max A \qquad A^* = [2, +\infty) \qquad \sup A = 2$$

and

$$\nexists \min A \qquad A_* = \emptyset \qquad \nexists \inf A = 2$$

If we are working on $\overline{\mathbb{R}}$ we have

$$\nexists \max A \qquad A^* = \emptyset \qquad \sup A = +\infty$$

and

$$\nexists \min A \qquad A_* = \emptyset \qquad \inf A = -\infty$$

Examples

Example 3.27. Consider the following subsets of $\mathbb{R}$

$$A = \{x \in \mathbb{R} : 0 \leq \max(x, -x) \leq 2\} \quad \text{and} \quad B = \{x \in \mathbb{R} : \min(x, -x) \leq -1\}$$

We have $A = [-2, 2]$ and $B = (-\infty, -1] \cup [1, +\infty)$. In $\mathbb{R}$, for A we have

$$\sup A = \max A = 2 \qquad \text{and} \qquad \inf A = \min A = -2$$

and A is a bounded set, while for B we have

$$\nexists \sup B = \max B \qquad \text{and} \qquad \nexists \inf B = \min B$$

and B is not a bounded set. In $\overline{\mathbb{R}}$, for A we have the same, while for B we have

$$\nexists \max B, \quad \sup B = +\infty, \quad \nexists \min B, \quad \inf B = -\infty$$

Supremum and infimum of a set

Theorem 3.28. (**Least Upper Bound Principle**). Each non-empty set $A \subseteq \mathbb{R}$ has supremum if it is bounded above and it is has infimum if it is bounded below.

Example 3.29. Let $A = (-2, 4)$ be a non-empty subset of $\mathbb{R}$. A is bounded (from above and below), therefore for Theorem 3.28, so A has supremum and infimum. We have

$$\sup A = 4 \quad \nexists \max A \quad \inf A = -2 \quad \nexists \min A$$

Example 3.30. Let A be a set defined by

$$A = \left\{ x \in \mathbb{R} : x = (n+1)^{-1}, \; n \in \mathbb{N} \right\}$$

We have

$$A = \left\{ 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots \right\}$$

therefore we can say that

$$\sup A = \max A = 1 \qquad \inf A = 0 \qquad \nexists \min A$$