STAT 538 Lecture 5

Convex sets

©Marina Meilă
mmp@stat.washington.edu

Reading: BV 2.1–2.3, 2.5. These notes are meant to guide the reading of BV, not to replace it.

- line segment $\{tx_2 + (1-t)x_2$ with $t \in [0,1]\}$
- affine combination $\sum_i t_i x_i$ with $\sum_i t_i = 1$
  Can also be written as $x_1 + \sum_{i=2:k} t_i(x_i - x_1)$ with any real $t_{2:k}$
- convex combination $\sum_i t_i x_i$ with $\sum_i t_i = 1$ and $t_i \geq 0$
- conic combination $\sum_i t_i x_i$ with $t_i \geq 0$
- cone $\mathcal{C}$: if $x \in \mathcal{C}$ then $tx \in \mathcal{C}$ for all $t > 0$
- convex set, affine set, cone, convex cone
- convex hull, affine hull, conic hull

Extreme points

Krein-Milman Theorem A bounded closed convex set (in $\mathbb{R}^d$) is the closed convex hull of its extreme points. [This theorem extends to spaces of functions too.]

Relative interior, (relative) boundary, (relative) closure

Examples of Convex sets

1. hyperplane $a^T(x - x_0) = 0$
2. half-space $a^T x - b \geq 0$
3. ball, ellipsoid
4. second-order cone $\{(x,t) \,|\, t \geq 0,\ ||x||_2 \leq t,\ x \in \mathbb{R}^d\}$
5. polyhedron = (bounded) intersection of m half-spaces
6. simplex = convex hull of $k+1$ affinely independent points (i.e $x_1 - x_{k+1}, x_2 - x_{k+1}, \ldots$ are linearly independent)
7. all symmetric matrices are an affine set and a convex set (unbounded)

8. all positive definite matrices are a convex cone
9. all stochastic matrices are a convex set
10. all doubly stochastic matrices are a convex set (what are its extreme points?)

Convex sets in probability

1. the parameter space of all normal distributions over $\mathbb{R}^d$ is a convex set
2. the (parameter) space of all discrete distributions over some countable space X
3. all distributions with a fixed set of marginals
4. the conditional distributions of a discrete joint
   Let $X, Y \in \Omega_X \times \Omega_Y$ with $\Omega_X| = m$, $|\Omega_Y| = n$ be two discrete random variables, and let Θ be the set of all probability distributions over $\Omega_X \times \Omega_Y$.
   That is, we define $P_\theta(X = i, Y = j) = \theta_{ij}$; then $\Theta \;=\; \{\theta = [\theta_{ij}]_{ij} \in [0,1]^{m\times n}, \sum_{ij} \theta_{ij} \;=\; 1\}$. Imagine θ to be rearranged as a vector of dimension mn:

   $$\mathrm{vec}(\theta) \;=\; [\,\theta_{11}\,\theta_{12}\ldots\theta_{21}\,\theta_{22}\,\ldots\theta_{mn}\,]^T \tag{1}$$

   We use the *linear-fractional (or projective) function* (BV page 41)

   $$f(z) \;=\; \frac{Az+b}{c^T z+d} \quad \mathrm{dom} f \;=\; \{z | c^T z + d > 0\} \tag{2}$$

   which maps a convex set into a convex set. Let now $z = \mathrm{vec}(\theta)$, $b = 0, d = 0$,

   $$A_{ij,i'j'} \;=\; \begin{cases} 1 & \text{if } j = j' = j_0, i' = i \\ 0 & \text{otherwise} \end{cases} \qquad c_{ij} \;=\; \begin{cases} 1 & \text{if } j = j_0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

   In other words, $c^T\mathrm{vec}(\theta) \;=\; \sum_i \theta_{ij_0} \;=\; P_\theta[Y \;=\; j_0]$, and row ij of A, multiplied by θ gives $\sum_{i',j'} A_{ij,i'j'}\theta_{i'j'} = \theta_{i,j_0} = P_\theta[X = i, Y = j_0]$.
   Hence, $f(\theta) = P_\theta[X = i|Y = j_0]$ for any $\theta \in \Theta$ with $P_\theta(Y = j_0) > 0$. This subset of Θ is also convex (but not closed), therefore we conclude that the set of all conditional probabilities given a fixed Y is a convex set.

5. all distributions over $\mathbb{R}$ with $E[X]$ in a convex set (in particular $E[X]$ fixed, $E[X] \geq a$)

Convex spaces of functions

1. polynomials of degree n; all polynomials
2. $\{g \,|\, g \geq f_0\}$ with f_0 fixed
3. $\{g \,|\, \int |g|^p < a\}$ with $p \geq 1$, $a \in (0, \infty)$ (the L_p balls)
4. all convex functions on set X

1 Separating and supporting hyperplanes

Theorem C, D convex sets, $C \cap D = \emptyset$. Then, there exist a hyperplane $a^T x - b = 0$ that separates C, D, i.e so that $a^T x \leq b$ for any $x \in C$ and $a^T x \geq b$ for any $x \in D$.

Strict separation = one of the inequalities is strict

Proposition C convex closed, $x_0 \notin C$. Then x_0 can be strictly separated from C.

Corollary: Theorem of alternatives (BV Example 2.21) The system of linear inequalities $Ax \prec b$, with $A \in \mathbb{R}^{m \times n}$ is infeasible iff the convex sets

$$C = \{b - Ax | x \in \mathbb{R}^n\} \quad \text{and} \quad D = \mathbb{R}^m_{++} = \{y \in \mathbb{R}^m | y \succ 0\} \qquad (4)$$

do not intersect.