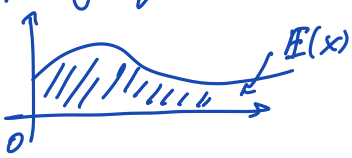
# Probabilistic bounds and applications to signal detection

(See more details in B&V 7.5)

Markov, Chebyshev, Chernoff

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a} \quad \text{for } X \in \mathbb{R}_+$$

Simple proof for distributions with densities:




$$\mathbb{E}(X) = \int_0^a x\, dp + \int_a^\infty x\, dp \geq 0 + \int_a^\infty a\, dp = a\,\mathbb{P}(X \geq a)$$

Chebyshev: $\mathbb{P}(|X-\mu| \geq 1) \leq \sigma^2$

$\mathbb{E}X = \mu, \quad \mathbb{E}(X-\mu) = \sigma^2$

Chernoff: gives more precise tail estimates for certain distributions (Bernoulli)

Goal: define more generic framework for convex optimization problems yielding such bounds

---

Key problem:

Let $X$ be a random variable on $S \subseteq \mathbb{R}^n$

$$\boxed{\begin{aligned} &\max \ \mathrm{Prob}(X \in C) \ -? \\ &\text{subject to } \ \mathbb{E} f_i(X) = a_i \end{aligned}}$$

$\leftarrow$ $f_i: \mathbb{R}^n \to \mathbb{R}$, $i = 1 \dots N$

e.g., moments of $X$

$$\mathbb{P}(X \in C) = \mathbb{E}(\mathbb{1}_C(X))$$

indicator function $\quad \mathbb{1}_C(z) = \begin{cases} 1 & \text{if } z \in C \\ 0 & \text{otherwise} \end{cases}$

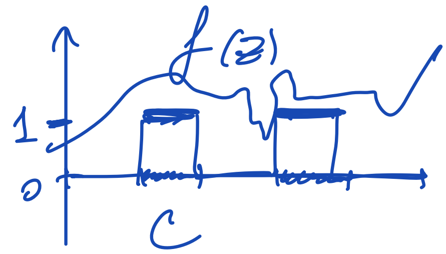
① Finite case: $X = x_i$ with probability $p_i$, $i = 1 \dots N$

$$\mathbb{E} f(X) = \sum_{i=1}^N p_i f(x_i)$$

$$\begin{bmatrix} \max \sum_{i: x_i \in C} p_i \\ \sum_{i=1}^N p_i f_j(x_i) = a_i \\ \sum p_i = 1 \end{bmatrix} \quad \text{LP}$$

② **General case:** consider $f(z) := \sum_{i=0}^{N} x_i f_i(z)$, $\mathbb{E} f_0(z) = 1$

Key observation: if $f(z) \geq \mathbb{1}_C(z)$, $\mathbb{E} f(z) \geq \mathbb{P}(X \in C)$




$$\boxed{\begin{aligned} &\min \quad x_0 + a_1 x_1 + \dots + a_n x_n \\ (*)\quad &\text{s.t.} \quad f(z) = \sum_{i=0}^{N} x_i f_i(z) \geq 1 \qquad z \in C \\ &\qquad\quad f(z) = \sum x_i f_i(z) \geq 0 \qquad z \in S \setminus C \end{aligned}}$$

$(x_0 \dots x_{N+1})$ are decision variables

$\Updownarrow$

or,

$$(*)\quad \boxed{\begin{aligned} &\min \quad \mathbb{E} f(X) \\ &\text{s.t.} \quad g_1(x) \leq 0 \\ &\qquad\quad g_2(x) \leq 0 \end{aligned}}$$

where

$\mathbb{E} f(z) = \langle a, x \rangle$, where $a = \begin{pmatrix} 1 \\ a_1 \\ \vdots \\ a_N \end{pmatrix}$

It is a convex problem:
$$\begin{cases} g_1(x) := 1 - \inf\limits_{z \in C} f(z) \leq 0 \\ g_2(x) := -\inf\limits_{z \in S \setminus C} f(z) \leq 0 \end{cases}$$

convex, as inf is a concave function

By key observation solution to the problem (*) upper bounds $\mathbb{P}(X \in C)$

**When is it easy to solve?**

ⓐ **Analytically:** some simple settings

Assume $S = \mathbb{R}_+$, $C = [a, \infty)$

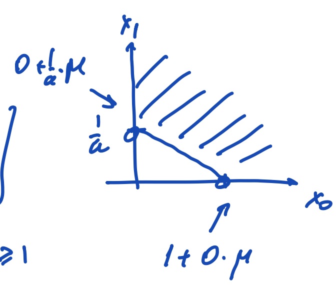
$\mathbb{E} X = \mu \leq a$

$(f_1(x) = x)$

$\min \quad x_0 + \mu x_1$

s.t. $\begin{bmatrix} x_0 + z x_1 \geq 1 & z \geq a \\ x_0 + z x_1 \geq 0 & \forall z \end{bmatrix}$

$x_0, x_1 \geq 0, \quad x_0 + a x_1 \geq 1$




$\min - \begin{cases} \frac{\mu}{a} & \text{if } \mu \leq a \\ 1 & \text{if } \mu \geq a \end{cases}$

$\swarrow$

So, $\mathbb{P}(X \geq a) \leq \frac{\mu}{a}$

ⓑ Case 2: with second moments

Assume $S = \mathbb{R}^n$ and

$\begin{bmatrix} \mathbb{E} X = \mu \in \mathbb{R}^n \\ \mathbb{E} \underbrace{X X^T}_{\Sigma_{ij} = \mathbb{E} X_i X_j} = \Sigma \in S^n \end{bmatrix}$ $\leftarrow$ $n$ functions with exp's $\mu_1 \dots \mu_n$

$\leftarrow n^2$ functions $f_{jk} = X_j X_k$

So, $f(z) = x_0 + \sum_{i=1}^{n} x_i z_i + \sum_{ij} x_{ij} z_i z_j$

or $f(z) = z^T P z + 2 q^T z + r$ ( $P \in Sym(n)$, $q \in \mathbb{R}^n$, $r \in \mathbb{R}$ are decision variables)

$\mathbb{E} f(X) = \mathbb{E}(X^T P X + 2q^T X + r) = \mathbb{E}(\operatorname{tr} P X X^T) + 2\mathbb{E} q^T x + r = \operatorname{tr}(\Sigma P) + 2q^T \mu + r.$

objective function: $\boxed{\min_{P,q,r} \mathbb{E} f(X)}$

Now, constraints:

- $f(z) \geq 0 \quad \forall z \iff \begin{bmatrix} P & q \\ q^T & r \end{bmatrix} \succeq 0$
- $f(z) \geq 1 \quad \forall z \in C.$

exterior of an open polytope

Further, let us assume $C = \mathbb{R}^n \setminus \mathcal{P}$, where $\mathcal{P} := \{z \mid a_i^T z < b_i, \quad i = 1 \dots k\}$

$z \in C$ means $\exists i: a_i^T z \geq b$ (or, $b_i - a_i^T x \leq 0$).

So, for any $i = 1 \dots k$, there is no $x$: $b_i - a_i^T x \leq 0$, but $f(z) < 1$ (or $f(z) - 1 = z^T P z + 2q^T z + r - 1 < 0$)

By Thm ✱ below, $\Longrightarrow \exists \tau_i: \begin{bmatrix} P & q \\ q^T & r-1 \end{bmatrix} \succeq \tau_i \begin{bmatrix} 0 & a_i/2 \\ a_i^T/2 & -b_i \end{bmatrix}.$

Thm ✱: Theorem of alternatives of a pair of quadratic inequalities;

**Thm** Suppose $\exists \hat{x}: \hat{x}^T A_2 \hat{x} + 2b_2^T \hat{x} + c_2 < 0$. Then

$\exists x$: $x^T A_1 x + 2b_1^T x + c_1 < 0$, $x^T A_2 x + 2b_2^T x + c_2 \leq 0$

$\iff \nexists \lambda \geq 0 \quad \begin{bmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{bmatrix} + \lambda \begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} \succeq 0.$

Proof: B&V § B2, B4

Remark: $0 \leq \begin{bmatrix} x \\ 1 \end{bmatrix}^T \left( \begin{bmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{bmatrix} + \lambda \begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix} \right) \begin{bmatrix} x \\ 1 \end{bmatrix} = x^T A_1 x + 2b_1^T x + c_1 + \lambda (x^T A_2 x + 2b_2^T x + c_2) < 0$

(weak alternative is obvious, together (1) and (2) lead to a contradiction)

As a conclusion, problem ✱ is of the form:

$$\begin{bmatrix} \min \operatorname{tr}(\Sigma P) + 2q^T a + r \\ \begin{bmatrix} P & q \\ q^T & r \end{bmatrix} \succeq 0 \\ \begin{bmatrix} P & q \\ q^T & r-1 \end{bmatrix} \succeq \tau_i \begin{bmatrix} 0 & a_i/2 \\ a_i^T/2 & -b_i \end{bmatrix} \\ \tau_i \geq 0 \end{bmatrix} = \alpha,$$

then $1-\alpha$ is a lower bound for the probability of a location inside the polytope

It is an SDP

Example $s \in \{s_1 \dots s_m\} \subseteq \mathbb{R}^n$ signal constellation

One of $s_i$'s is transmitted via a noisy channel

received $x = s + v$ ← random noise

$\mathbb{E}v = 0 \quad \mathbb{E}vv^T = \sigma^2 I$

Minimum distance estimator: $s_k$ closest to $x$

Prob (correct detection)?

It is given by a polytope:

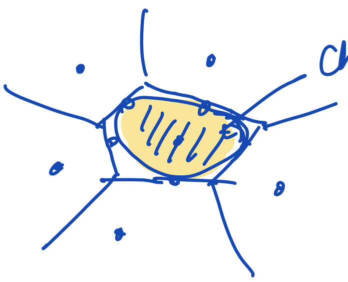
$$\|x - s_k\|_2 < \|x - s_j\|_2 \qquad j \neq k$$

$$\|v\|_2^2 \le \|v + s_k - s_j\|_2^2$$

$$\vdots$$

$$2 \langle s_j - s_k, v + s_k \rangle \le \|s_j\|^2 - \|s_k\|^2 \quad \text{for each } j \neq k$$

Voronoi region $V_k$



Chebyshev estimate lower bound for prob.

(probability of correct detection of each of the signals depending on $\sigma$)

Fig 7.6 Boid & Van

$x^T P x + 2q^T x + r \ge 1$
with optimal $P, q, r$
define yellow ellipsoid