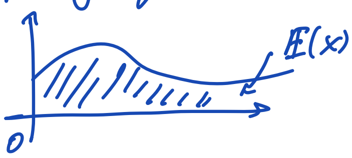
Probabilistic bounds and applications to signal detection

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 $S \subseteq \mathbb{R}^n$

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

$\leftarrow$ e.g., moments of X $\quad f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1 \ldots N$

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

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

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

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

$$\left[\begin{array}{l} \max \sum_{i : x_i \in C} p_i \\ \sum_{i=1}^N p_i \, f_j(x_i) = a_i \\ \sum p_i = 1 \end{array} \right] \quad \text{LP}$$

② <u>General case</u>: consider $f(z) := \sum_{i=0}^{N} x_i f_i(z)$

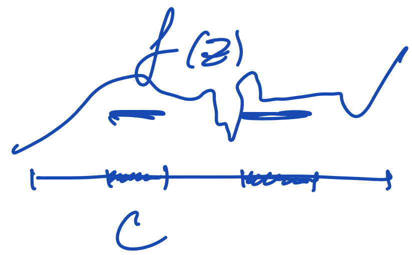
$\mathbb{E} f_0(X) := 1$

If $f(z) \geq \mathbb{1}_C(z)$, $\underbrace{\mathbb{E} f(z)}_{\langle a, x\rangle} \geq \mathbb{P}(X \in C)$ (*)

$\langle a, x \rangle$, where $a = \begin{pmatrix} 1 \\ a_1 \\ \vdots \\ a_N \end{pmatrix}$

So, $$\left[\begin{array}{lll} \min & x_0 + a_1 x_1 + \dots + a_n x_n & \\ \text{s.t.} & f(z) = \sum_{i=0}^{N} x_i f_i(z) \geq 1 & z \in C \\ & f(z) = \sum x_i f_i(z) \geq 0 & z \in S \setminus C \end{array}\right]$$

$x \in \mathbb{R}^{N+1}$



It is convex
$$g_1(x) = 1 - \inf_{z \in C} f(z) \leq 0$$
$$g_2(x) = -\inf_{z \in S \setminus C} f(z) \leq 0$$
convex, or "infinitely many linear conditions"

When is it easy to solve?

ⓐ Analytically: some simple settings

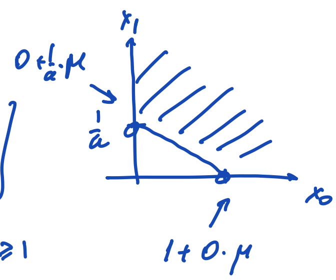
Assume $S = \mathbb{R}_+$, $C = [a, \infty)$

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

$(f_1(x) = x)$

$$\min \; x_0 + \mu x_1$$
$$\text{s.t.} \left[\begin{array}{ll} x_0 + z x_1 \geq 1 & z \geq a \\ x_0 + z x_1 \geq 0 & \forall z \end{array}\right]$$

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



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

from (*)
$$\mathbb{P}(X \geq a) \leq \frac{\mu}{a}$$

ⓑ $S = \mathbb{R}^n$

$$\left[\begin{array}{l} \mathbb{E} X = \mu \in \mathbb{R}^n \\ \mathbb{E} X X^T = \Sigma \in S^n \end{array}\right]$$
$\leftarrow$ n functions with exp's $\mu_1 \dots \mu_n$

$\leftarrow$ n^2 functions $f_{jk} = x_j x_k$

So,

$$f(z) = x_0 + \sum_{i=1}^{n} x_i z_i + \sum_{j,k=1}^{n} z_j z_k x_{jk}$$

$$f(z) = z^T P z + 2q^T z + r \longleftarrow \quad P, q, r \text{ contain } x\text{'s}$$

$$\mathbb{E} f(X) = \mathbb{E}\left(X^T P X + 2q^T X + r\right)$$

$$= \mathbb{E}\, \mathrm{tr}(PXX^T) + 2\mathbb{E} q^T X + r = \mathrm{tr}(\Sigma P) + 2q^T \mu + r$$

$$f(z) \geq 0 \ \forall z \iff \begin{bmatrix} P & q \\ q^T & r \end{bmatrix} \succeq 0$$

Let $C = \mathbb{R}^n \setminus \mathcal{P}$ $\quad \mathcal{P} := \{z \mid a_i^T z < b_i, \ i = 1 \dots k\}$ open polytope

$f(z) \geq 1$ for $z \in C$:

$\circledast$ If $a_i^T z \geq b_i$ for some i, then $z^T P z + 2q^T z + r \geq 1$

$\circledast$ is equivalent to:

$$\exists \tau_1 \dots \tau_n \geq 0: \quad \begin{bmatrix} P & q \\ q^T & r-1 \end{bmatrix} \succeq \tau_i \begin{bmatrix} 0 & a_i/2 \\ a_i^T/2 & -b_i \end{bmatrix} \quad i = 1 \dots n$$

Why? 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

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

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

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)

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

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

SDP

Example $s \in \{s_1 \ldots s_m\} \subseteq \mathbb{R}^n$ 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:

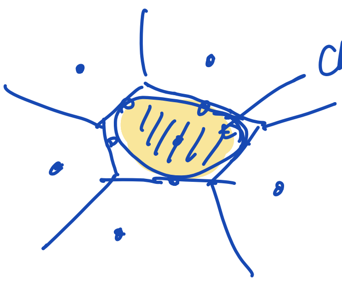
$$\|x - s_k\|_2 < \|x - s_j\|_2 \qquad j \neq k$$

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

$$\vdots$$

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

Voronoi region V_k



Chebyshev estimate lower bound for prob.

(probability of correct detection of each of the signals depending on σ)

Fig 8.6 Boid & Van

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