EE 227A: Convex Optimization and Applications **April 17, 2012**

Lecture 21: **Moment Inequalities**

Lecturer: Laurent El Ghaoui

Reading assignment: Section §7.4 of BV, and the Chapter 3 of the book on Robust Optimization[1]

21.1 Chance Linear Programming

See the Chapter mentioned above.

21.2 Bounds on Probabilities

21.2.1 Problem statement

We consider a random variable x with distribution π, which is only known to belong to a class of distributions Π, and seek a bound on the probability of a given set $\mathcal{C}$, that is, a lower bound on

$$\inf_{\pi \in \Pi} \mathbf{Prob}\,\mathcal{C}. \tag{21.1}$$

Alternatively, we may seek an upper bound on the quantity obtained by replacing "inf" with "sup" in the above. Both problem are equivalent, in the sense that replacing $\mathcal{C}$ by its complement in one problem leads to the other.

We assume that Π is a class of distributions with given mean and covariance matrix:

$$\Pi = \left\{ \pi \in \Pi_0 \;:\; \mathbf{E}_\pi x = \hat{x}, \;\; \mathbf{E}_\pi (x - \hat{x})(x - \hat{x})^T = \Sigma \right\},$$

where Π_0 is the set distributions on $\mathbf{R}^n$, $\hat{x} \in \mathbf{R}^n$, $\Sigma \in \mathcal{S}^n_{++}$ are given, and $\mathbf{E}_\pi$ denotes the expectation operator with respect to the distribution π.

Problems involving bounds on probabilities arise in many situations. For example, we may interested in *yield maximization,* which involves the function $Y(y) := \mathbf{Prob}(y + x \in S)$, where y is a vector of design parameters, x represents additive implementation noise, and S is a a subset of allowable designs.

21.2.2 Dual problem

We can formulate problem (21.1) as an infinite dimensional linear programming problem:

$$p^* := \inf_{\pi(\cdot) \geq 0} \int \mathbf{1}_{\mathcal{C}}(x)\pi(x)dx \;:\; \int \begin{pmatrix} x \\ 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}^T \pi(x)dx = \Gamma, \tag{21.2}$$

[1] A. Ben Tal, L. El Ghaoui, A. Nemirovski, *Robust Optimization.*

where $\mathbf{1}_{\mathcal{C}}$ is the location function of $\mathcal{C}$ (with value 1 on $\mathcal{C}$, and 0 elsewhere), and

$$\Gamma = \begin{pmatrix} \Sigma + xx^T & x \\ x^T & 1 \end{pmatrix} \succ 0.$$

The problem is linear in the sense that it involves an objective that is linear (in the variable π), affine equality constraints, and sign constraints. Of course, this is not an LP in the classical sense, as the variable is infinite-dimensional.

Using duality, we can transform the problem into one with infinitely many constraints, and finitely many variables. To do this, we first obtain a weak duality result, using the Lagrange functional[2]

$$\mathcal{L}(\pi, M) = \inf_{\pi(\cdot)\geq 0} \int \mathbf{1}_{\mathcal{C}}(x)\pi(x)dx + \langle M, \Gamma - \int \begin{pmatrix} x \\ 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}^T \pi(x)dx \rangle.$$

We check that this Lagrangian "works", that is, we have the minimax representation

$$p^* := \min_{\pi(\cdot)\geq 0} \max_{M\in\mathcal{S}^n} \mathcal{L}(\pi, M).$$

By weak duality, we have $d^* \leq p^*$, with

$$d^* := \max_{M\in\mathcal{S}^n} \min_{\pi(\cdot)\geq 0} \mathcal{L}(\pi, M).$$

The dual function is

$$g(M) := \min_{\pi(\cdot)\geq 0} \mathcal{L}(\pi, M) = \langle M, \Gamma \rangle + \min_{\pi(\cdot)\geq 0} \langle \pi, \mathbf{1}_{\mathcal{C}} - q_M \rangle,$$

where q_M is the quadratic function with values

$$q_M(x) := \langle M, \begin{pmatrix} x \\ 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}^T \rangle = \begin{pmatrix} x \\ 1 \end{pmatrix}^T M \begin{pmatrix} x \\ 1 \end{pmatrix},$$

and we define the scalar product between two measures π, h as

$$\langle \pi, h \rangle := \int \pi(x)h(x)dx.$$

It is easy to show that, for any function h:

$$\min_{\pi(\cdot)\geq 0} \langle \pi, h \rangle = \begin{cases} 0 & \text{if } h(x) \geq 0 \text{ for every } x \in \mathbf{R}^n, \\ -\infty & \text{otherwise.} \end{cases}$$

[2]We say "functional" as the Lagrangian's input variables includes π, which is a function, more precisely a measure.

We obtain

$$g(M) = \begin{cases} \langle M, \Gamma \rangle & \text{if } \mathbf{1}_{\mathcal{C}}(x) \geq q_M(x) \text{ for every } x \in \mathbf{R}^n, \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem reads

$$\sup_{M=M^T} g(M) = \sup_{M=M^T} \langle M, \Gamma \rangle \ : \ \begin{array}{l} \forall x \in \mathbf{R}^n, q_M(x) \leq 1, \\ \forall\, x \notin \mathcal{C}, \ \ q_M(x) \leq 0. \end{array}$$

The first constraint is equivalent to the semidefinite constraint $M \preceq J$, where J is a matrix with all zeros, except a 1 in the lower-right element.

Further reductions. In some cases, the dual problem can be expressed exactly as a semidefinite program. Consider the case when $\mathcal{C}$ is defined by a *single* (possibly non-convex) quadratic inequality:

$$\mathcal{C} = \left\{ x \ : \ q(x) := \begin{pmatrix} x \\ 1 \end{pmatrix}^T Q \begin{pmatrix} x \\ 1 \end{pmatrix} > 0 \right\},$$

with $Q = Q^T$ given.

Then, using the $\mathcal{S}$-lemma (see BV, §B.2) the condition

$$\forall\, x, \ \ q(x) \leq 0, \ \ q_M(x) \leq 0$$

is equivalent to the existence of $\tau \geq 0$ such that $M \preceq \tau Q$.

The dual problem now reads

$$d^* = \sup_{M=M^T,\, \tau \geq 0} \langle M, \Gamma \rangle \ : \ J \succeq M, \ \ M \preceq \tau Q,$$

which is an SDP. It turns out that it can be further simplified greatly, to a single-variable convex problem. Precisely, we have

$$1 - d^* = \min_{\tau \geq 0} \ \lambda_{\max}[\Gamma^{1/2}(J - \tau Q)\Gamma^{1/2}]_+,$$

where X_+ is the matrix obtained from the symmetric matrix X by replacing the negative eigenvalues by 0.

Strong duality. It can be shown that if $\Gamma \succ 0$, then strong duality holds.

Chebyschev and Markov inequalities. Chebyschev and Markov inequalities can be derived from the above, as special cases.