ISyE 6664 Stochastic Optimization

Fall 2012

Assignment 2

Issued: September 6, 2012

Due: September 18, 2012

Problem 1
Consider the newsvendor problem in which the demand D has a known probability distribution with cumulative distribution function F. No other assumptions are made regarding F (whether F has a density, or whether F is a discrete cdf, etc.). Show that, in general, the set of optimal order quantities is given by

$$F^{-1}\left(\frac{r-c}{r-s}\right)$$

where $F^{-1}(p)$ denotes the set of p-quantiles of F, that is

$$F^{-1}(p) \quad \equiv \quad \{x \in \mathbb{R} \,:\, F(y) \leq p \;\forall\, y < x \text{ and } F(y) \geq p \;\forall\, y \geq x\}$$

Problem 2
Consider a random variable D with cumulative distribution function F. Let $D_1, \ldots, D_n$ be an i.i.d. random sample of D with empirical cumulative distribution function $\hat{H}_n$. Consider any $p \in (0,1)$, and let $F^{-1}(p) = [a,b]$, and $\hat{H}_n^{-1}(p) = [a_n, b_n]$.

1. For a set $A \subset \mathbb{R}$ and $x \in \mathbb{R}$, let $d(x,A) := \inf\{|x-y| \,:\, y \in A\}$ denote the distance between x and A. For two sets $A, B \subset \mathbb{R}$, let $e(B,A) := \sup\{d(x,A) \,:\, x \in B\}$ denote the excess of B over A. (Note that e is not symmetric; $\max\{e(A,B), e(B,A)\}$ is a metric.) Use the strong law of large numbers to show that $\mathbb{P}[e([a_n,b_n],[a,b]) \to 0 \text{ as } n \to \infty] = 1$, that is, $[a_n, b_n]$ moves close to $[a,b]$ as $n \to \infty$.

2. Show with a counterexample that it does not always hold that $\mathbb{P}[e([a,b],[a_n,b_n]) \to 0 \text{ as } n \to \infty] = 1$, that is, the empirical quantile may not be close to the entire true quantile for large n.

3. Show that $\mathbb{P}[\min\{|a_n - a|, |b_n - b|\} \to 0 \text{ as } n \to \infty] = 1$, that is, at least one of the empirical left quantile or empirical right quantile is close to the true left quantile or true right quantile respectively for large n.

4. Show with a counterexample that it does not always hold that $\mathbb{P}[\max\{|a_n - a|, |b_n - b|\} \to 0 \text{ as } n \to \infty] = 1$, that is, the empirical quantile may not be close to the entire true quantile for large n.

Problem 3

Randomization does not help for static optimization problems: Suppose we want to maximize $g(x)$ over all $x \in \mathcal{X}$. That is, the optimization problem is

$$\sup_{x \in \mathcal{X}} g(x)$$

(The following also holds if $g(x) \equiv E[G(x,\omega)]$.) A deterministic decision chooses one particular decision $x \in \mathcal{X}$. A randomized decision chooses a probability distribution P on $\mathcal{X}$ (we can choose a sufficiently rich σ-field on $\mathcal{X}$), and then a decision $x \in \mathcal{X}$ is generated according to probability distribution P. The objective value of such a randomized decision P is $E_P[g(X)]$ (where $g(X)$ is a random variable if the σ-field on $\mathcal{X}$ is sufficiently rich).

Show that randomization does not help for such a static optimization problem, unless $\sup_{x \in \mathcal{X}} g(x) = \infty$. (There are dynamic optimization problems for which randomization does help. Furthermore, there are worst-case type problems for which randomization may *appear* to help.) That is, show the following:

1. For any probability distribution P on $\mathcal{X}$,

   $$E_P[g(X)] \quad \leq \quad \sup_{x \in \mathcal{X}} g(x)$$

   That is,

   $$\sup_{P \in \mathcal{P}} E_P[g(X)] \quad = \quad \sup_{x \in \mathcal{X}} g(x)$$

   where $\mathcal{P}$ denotes the set of probability distributions on $\mathcal{X}$.

2. If $g^* \equiv \sup_{x \in \mathcal{X}} g(x) < \infty$ and there exists no $x \in \mathcal{X}$ such that $g(x) = g^*$, that is, there exists no deterministic decision $x \in \mathcal{X}$ that attains the supremum, then there exists no $P \in \mathcal{P}$ such that $E_P[g(X)] = g^*$, that is, there exists no randomized decision $P \in \mathcal{P}$ that attains the supremum.

3. Suppose we have a static optimization problem where $g(x) < \infty$ for all $x \in \mathcal{X}$, but $\sup_{x \in \mathcal{X}} g(x) = \infty$. Show how to choose a probability distribution P such that $E_P[g(X)] = \infty$. That is, there exists no deterministic decision $x \in \mathcal{X}$ that attains the supremum, but there exists a randomized decision $P \in \mathcal{P}$ that attains the supremum.

Problem 4

Randomization may *appear* to help for worst-case problems: Consider a function $G(x,\omega)$, with decision $x \in \mathcal{X}$, and $\omega \in \Omega$ unknown at the time the decision has to be made. Suppose we want to maximize $G(x,\omega)$ for the worst possible outcome $\omega \in \Omega$. That is, the optimization problem is

$$\sup_{x \in \mathcal{X}} \inf_{\omega \in \Omega} G(x,\omega)$$

A deterministic decision chooses one particular decision $x \in \mathcal{X}$. A randomized decision chooses a probability distribution P on $\mathcal{X}$, and then a decision $x \in \mathcal{X}$ is generated according to probability distribution P. In case such a randomized decision P is chosen, the objective function takes the expected value $E_P[\cdot]$ over all decisions X, because X is now a random variable distributed according to P. Exactly how this expected value is taken has important consequences, because for any probability distribution P,

$$\inf_{\omega \in \Omega} E_P \left[G(X,\omega)\right] \quad \geq \quad E_P \left[\inf_{\omega \in \Omega} G(X,\omega)\right]$$

and thus

$$\sup_{P \in \mathcal{P}} \inf_{\omega \in \Omega} E_P \left[G(X,\omega)\right] \quad \geq \quad \sup_{P \in \mathcal{P}} \; E_P \left[\inf_{\omega \in \Omega} G(X,\omega)\right]$$

To interpret the difference, it is useful to think of a decision maker that chooses P to maximize $G(x,\omega)$ and an adversary that chooses ω to minimize $G(x,\omega)$. First, the decision maker chooses P. If thereafter the decision x is generated according to P, and then the adversary gets to observe x and then chooses ω (such an adversary is sometimes called an omniscient adversary), then the outcome of the decision process is given by

$$\sup_{P \in \mathcal{P}} \; E_P \left[\inf_{\omega \in \Omega} G(X,\omega)\right]$$

Let $g_1(x) \equiv \inf_{\omega \in \Omega} G(x,\omega)$ for all x. Thus, the outcome of the decision process is

$$\sup_{P \in \mathcal{P}} \; E_P \left[g_1(X)\right]$$

We know that

$$\sup_{P \in \mathcal{P}} \; E_P \left[\inf_{\omega \in \Omega} G(X,\omega)\right] \quad = \quad \sup_{x \in \mathcal{X}} \; \inf_{\omega \in \Omega} \; G(x,\omega)$$

and if $\sup_{x \in \mathcal{X}} g_1(x) < \infty$, then randomization does not help, and the decision maker is just as well off making a deterministic decision x, and hence we are back at the problem

$$\sup_{x \in \mathcal{X}} \; \inf_{\omega \in \Omega} \; G(x,\omega)$$

However, if we imagine that the decision maker chooses P, and thereafter the adversary has to choose ω, knowing the decision maker's choice of P but not the resulting decision x (such an adversary is sometimes called an oblivious adversary), then the outcome is given by

$$\sup_{P \in \mathcal{P}} \; \inf_{\omega \in \Omega} E_P \left[G(X,\omega)\right]$$

In this case we can exploit the property that

$$\sup_{P \in \mathcal{P}} \; \inf_{\omega \in \Omega} E_P \left[G(X,\omega)\right] \quad \geq \quad \sup_{P \in \mathcal{P}} \; E_P \left[\inf_{\omega \in \Omega} G(X,\omega)\right]$$

to make randomized decisions appear better than deterministic decisions. Let

$$g^* \quad \equiv \quad \sup_{x \in \mathcal{X}} \; \inf_{\omega \in \Omega} \; G(x,\omega)$$

and assume that $g^* < \infty$. (If $g^* = \infty$ then we can choose a probability distribution P on $\mathcal{X}$ such that $E_P\left[\inf_{\omega\in\Omega} G(X,\omega)\right] = \infty$, and thus the expected value of the outcome is ∞ for both the case with the omniscient adversary and the case with the oblivious adversary.) For ease of comparison, suppose that x^* is an optimal deterministic decision, that is, $G(x^*,\omega) \geq g^*$ for all $\omega \in \Omega$. For many problems, a decision x^* is good for some values of ω and not good for other values of ω. Suppose we have a set $\{x_1,\ldots,x_n\} \subset \mathcal{X}$ of n optimal deterministic decisions, each of which is good on some subset of Ω and reasonable on the rest of Ω. (Obviously, there are many examples of problems for which this holds—think about this a little.) Specifically, suppose Ω can be partitioned into n subsets, $\Omega = \cup_{i=1}^n \Omega_i$, with $\Omega_i \cap \Omega_j = \varnothing$ for all $i \neq j$, such that $G(x_i,\omega) \geq g^*$ for all $\omega \in \Omega$ and

$$g_i^* \quad \equiv \quad \inf_{\omega\in\Omega_i} G(x_i,\omega) \quad > \quad g^*$$

(that is, x_i is good on subset Ω_i) for each $i \in \{1,\ldots,n\}$. (Note that, in general, for any x,

$$\inf_{\omega\in\Omega_i} G(x,\omega) \quad \geq \quad \inf_{\omega\in\Omega} G(x,\omega)$$

because $\Omega_i \subset \Omega$.) Now show how to construct a randomized decision $\tilde{P}$ such that

$$\inf_{\omega\in\Omega} E_{\tilde{P}}\left[G(X,\omega)\right] \quad > \quad \inf_{\omega\in\Omega} G(x^*,\omega) \quad = \quad g^*$$

that is, it seems as if randomization helps. (This approach is popular in the computer science community for constructing "randomized algorithms".) Do you think the randomized decision is really better than optimal deterministic decision x^*? Why or why not?