Convergence of stochastic average approximation for stochastic optimization problems with mixed expectation and per-scenario constraints ☆

Mihai Anitescu[a,1], John R. Birge[b,2]

[a] Argonne National Laboratory, Mathematics and Computer Science Division, 9700 S. Cass Avenue, Argonne, IL 60439, USA, anitescu@mcs.anl.gov
[b] The University of Chicago Booth School of Business, 5807 S Woodlawn Avenue, Chicago, IL 60637, USA, John.Birge@chicagogsb.edu

Abstract

We present a framework for ensuring convergence of sample average approximations to stochastic optimization problems that include expectation constraints in addition to per-scenario constraints.

Key words: Sample average approximation, stochastic optimization, expectation constraints

1. Introduction

Stochastic optimization problems with mixed expectations and per-scenario constraints (SOESC) are ubiquitous in applications. As an example problem, consider an independent system operator (ISO) of an electric power network. In response to bids from a set of generators, the ISO agrees to purchase quantities of electricity $x \in \mathbb{R}^n$ for a future period at prices $\pi_0 \in \mathbb{R}^n$. In order to ensure participation in the market and sufficient supply, the prices π should represent forward prices as the expectation (under a risk-neutral measure) of future spot prices $\pi(\omega)$; so, $\pi_0 - E_\omega[\pi(\omega)] = z^+ - z^-$, where $\pi(\omega)$ satisfies an equilibrium condition for the future market under a random outcome ω and z^+ and z^- are non-negative variables satisfying a complementarity condition on the initial purchase quantities. Other forms of expectation constraints arise from risk considerations, where, for example, $r(x, \omega)$ is a risk function, such as excess loss, associated with outcome ω that must be compensated in expectation with allocated capital, x_0, as $x_0 - E_\omega[r(x, \omega)] = 0$. Expectation constraints also arise when the non-anticipativity of first-stage decisions is given explicitly as $x - E_\omega[x(\omega)] = 0$. In the following, we represent these constraints generally as $E_\omega[\psi(x, y(\omega), \omega)] = 0$, where x represents the first-stage, or upper-level, decisions and $y(\omega)$ represents the second-stage, or lower-level, decision. In addition, the problems of the generators have per-scenario constraints, resulting in a SOESC problem.

Other examples of applications formulated as SOESC problems include portfolio optimization with conditional value-at-risk objectives and constraints [3] and stochastic receding horizon control of constrained systems [6].

This paper is concerned with convergence of sample average approximation (SAA) approaches for SOESC. While convergence of SAA has been amply analyzed for per-scenario-only constraints [1, 7] and for expectations-only constraints [2, 5, 8], a similar analysis for SOESC is lacking. This paper takes an initial step toward filling that gap.

☆Preprint ANL/MCS P1562-1108

[1] Mihai Anitescu was supported by the Department of Energy, Contract No. DE-AC02-06CH11357.

[2] John Birge was supported by The University of Chicago Booth School of Business.

2. Formulating SOESC as mixed nonlinear equation – optimization problems

SOESC can be formulated as

$$\begin{array}{rl} \min\limits_{x \in K_x, y(\omega) \in K_y} & E_\omega\left[\phi(x, y(\omega), \omega)\right] \\ \text{such that} & 0 = E_\omega\left[\psi(x, y(\omega), \omega)\right]. \\ & 0 = \Gamma(x, y(\omega), \omega),\ \forall \omega \in \Omega. \end{array} \tag{1}$$

Here $\phi : \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}$, $\psi : \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}^p$, $\Gamma : \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}^q$, are differentiable functions, and K_x, K_y are closed sets. We note that both inequality and complementarity constraints can be formulated in this fashion, by an appropriate choice of K_x and K_y.

We define the partial Lagrangian, $\mathcal{L}^p(x, y(\omega), \omega, \lambda) =$

$$\phi(x, y(\omega), \omega) + \lambda^T \psi(x, y(\omega), \omega).$$

With it, we define the following mixed minimization–nonlinear equation problem:

$$\begin{array}{l} 0 = E_\omega\left[\psi(x, y(\omega), \omega)\right]; (x, \{y(\omega)\}) = \\ \underset{x \in K_x, y(\omega) \in K_y}{\operatorname{argmin}} \quad E_\omega\left[\mathcal{L}^p\left(x, y(\omega), \omega, \lambda\right)\right] \\ \text{s.t. } \Gamma(x, y(\omega), \omega) = 0, \forall \omega \in \Omega. \end{array} \tag{2}$$

In some circumstances, the formulation (2), which originates in a Lagrangian relaxation, is equivalent to the formulation (1). An example of sufficient conditions for this to occur is presented in Section 4. Should such an equivalence be determined, then $y(\omega)$ satisfies, based on the reversibility principle, the following relationship:

$$\tilde{y}(\omega, \lambda, x) = \begin{array}{l} \underset{y(\omega) \in K_y}{\operatorname{argmin}} \mathcal{L}^p(x, y(\omega), \omega, \lambda) \\ \text{s.t. } \Gamma(x, y(\omega), \omega) = 0, \forall \omega \in \Omega, \end{array} \tag{3}$$

provided that, for each ω, λ, x in the neighborhoods of interest, the solution exists and is unique.

With this framework, the solution of the minimization-nonlinear equation problem (2) can be restated as

$$\begin{array}{l} E_\omega\left[\psi(x, \tilde{y}(\omega, x, \lambda), \omega)\right] = 0 \\ x = \underset{x \in K_x}{\operatorname{argmin}}\ E_\omega\left[\mathcal{L}^p(x, \tilde{y}(\omega, x, \lambda), \omega, \lambda)\right]. \end{array} \tag{4}$$

Because y can be determined *independently* for different ω, λ, x and since, once those are determined, the remaining variables are deterministic and finite dimensional, the problem (4) is approachable by SAA.

What we must now show is that, should y be determined by SAA, then the resulting approximation of the data functions for (4) approximates uniformly the solution set for the minimization problem and nonlinear equations, respectively, with probability one.

3. Convergence of approximations of mixed nonlinear equations – optimization problems

Assumption We assume that the functions f^N, g^N, f, g appearing in this section are continuous in (x, λ).

Notation All sequences of functions indexed by ^{N} depend on the stochastic parameter ω, but we do not explicitly represent that dependence (e.g. $f^N(x, \lambda, \omega) \sim f^N(x, \lambda)$).

The various convergence concepts used here can be found in [7]. The following concept is central to our work:

Definition [7, 6.86]: Uniform convergence with probability one (UC w.p.1). $f^N \to f$ w.p. 1 if, for almost all ω and for any ϵ, there exists $N^*(\epsilon, \omega)$ such that

$$\sup_{x \in X, \lambda \in \Lambda} \left| f^N(x, \lambda) - f(x, \lambda) \right| \le \epsilon.$$

An important consequence of UC w.p. 1 is that, given a sequence $(x^N, \lambda^N) \to (x^*, \lambda^*)$, we have that $f^N(x^N, \lambda^N) \to f(x^*, \lambda^*)$ w.p.1.

The first few results extend results in [7] to the case of functions depending on a parameter.

Lemma 1. *Let $x \in X$, a compact set. Let $v(\lambda) = \min_{x \in X} f(x, \lambda)$ and let $S(\lambda)$ be the solution set of the same problem; then (a) $v(\lambda)$ is continuous; and (b) the following holds*

$$\limsup_{\lambda \to \lambda^*} \sup_{x \in S(\lambda)} \inf_{x^* \in S(\lambda^*)} |x - x^*| = 0.$$

Proof (a) Let $\lambda^n \to \lambda^*$. Let now x^n be a sequence satisfying $f(x^n, \lambda^n) = v(\lambda^n)$, and let x^* be such that $f(x^*, \lambda^*) = v(\lambda^*)$. We then have that

$$v(\lambda^n) = f(x^n, \lambda^n) \le f(x^*, \lambda^n) \overset{\lambda \to \lambda^*}{\to} v(\lambda^*)$$

and thus $\limsup v(\lambda^n) \le v(\lambda^*)$. In addition,

$$v(\lambda^*) = f(x^*, \lambda^*) \le f(x^n, \lambda^*) \\ = v(\lambda) + f(x^n, \lambda^*) - f(x^n, \lambda^n).$$

By continuity of f, this implies $v(\lambda^*) \le \liminf v(\lambda^n)$ which proves the claim of part (a).

(b) Assume that this claim is not true. That is, there exist $\epsilon > 0$ and a subsequence $\lambda^n \to \lambda^*$, such that $\sup_{x^n \in S(\lambda^n)} \inf_{x \in S(\lambda^*)} \|x^n - x^*\| \ge \epsilon$. Since $x^n \in X$ and X is compact, it has an accumulation point $\hat{x}$. Because of the above relationship, we must have that $\hat{x} \notin S(\lambda^*)$, and thus $f(\hat{x}, \lambda^*) > v(\lambda^*)$. From an argument similar to the one in part (a), and using the convergence of x^n to $\hat{x}$, we obtain that $f(\hat{x}, \lambda^*) = \lim f(x^n, \lambda^n) = \lim v(\lambda^n) = v(\lambda^*)$ and, thus a contradiction. □

Lemma 2. *Suppose that there exists a compact set C such that*

(i) $\emptyset \ne S(\lambda) \subset C$;

(ii) $f(x, \lambda)$ *is finite valued and continuous on* $C \times \Lambda$, *where* Λ *is a compact set;*

(iii) $\hat{f}^N(x, \lambda) \to f(x, \lambda)$ *w.p. 1 as* $N \to \infty$, *uniformly for* $x \in C$.

Then, $D(S^N(\lambda), S(\lambda)) \to 0$ *w.p. 1, uniformly in* λ. *Here we denote by* $S^N(\lambda)$ *is the solution set of* $\min_{x \in X} f^N(x, \lambda)$, *by* $D(\cdot, \cdot)$ *the distance between two sets, and by* $d(\cdot, \cdot)$ *the distance between a point and a set.*

Proof We assume, WLOG, that $f^N \to f$ for all ω, uniformly in (x, λ).

Assume that $D(S^N(\lambda), S)$ does not converge to 0 w.p. 1 uniformly. Then, there exist an ω and an ϵ and $N_k \to \infty$ such that, for any k, there exists $x^{N_k} \in S^{N_k}(\lambda^{N_k})$ such that

$$\inf_{x \in S(\lambda^{N^k})} \left| x^{N_k} - x \right| \ge \epsilon, \forall k. \qquad (5)$$

Since Λ and C are compact, we can extract a subsequence N_{k_j} such that $x^{N_{k_j}}$ and $\lambda^{N_{k_j}}$ are convergent to x^* and λ^*. To simplify the notation, we relabel the sequence N_{k_j} as N.

From Lemma 1, for N sufficiently large, we must have that, for λ^N is sufficiently close to λ^*, $D(S(\lambda^N), S(\lambda^*)) \le \frac{\epsilon}{4}$.

In turn this, in addition to (5), implies that, for N sufficiently large, we have

$$\inf_{x \in S(\lambda^*)} \left| x^N - x \right| \ge \frac{3\epsilon}{4} \; \forall N,$$

which then implies that $f(x^*, \lambda^*) > v(\lambda^*)$.

Since $x^N \in S^N(\lambda^N)$, we have that $f^N(x^N, \lambda^N) < f^N(\tilde{x}, \lambda^N)$, where $\tilde{x} \in S(\lambda^*)$. From the uniform convergence of f^N, we have that

$$\begin{aligned} f(x^*, \lambda^*) &= \lim f^N(x^N, \lambda^N) \\ &\le \lim f^N(\tilde{x}, \lambda^N) \\ &= f(\tilde{x}, \lambda^*) = v(\lambda^*), \end{aligned}$$

which is a contradiction. □

We now investigate the stability of systems of nonlinear equations.

Lemma 3. *Let C be a compact set. We define* $F^N = \left\{x | f^N(x) = 0\right\} \cap C$ *and* $F = \{x | f(x) = 0\} \cap C$. *If* $f^N \to f$ *UC w. p. 1, then* $D(F^N, F^*) \to 0$ *w.p. 1.*

Proof We prove this statement by contrapositive. There then exists a sequence (which we denote WLOG the same as the original sequence), such that there exists $x^N \in C$ that satisfies $F^N(x^N) = 0$, and $d(x^N, F) > \epsilon$. Since C is compact, we can assume (after extracting a subsequence and relabeling) that $x^N \to x^*$. Since $d(x^N, F) > \epsilon$, it follows that $d(x^*, F) > \epsilon$; but f^N converges uniformly and, therefore, $0 = f^N(x^N) \to f(x^*)$, which is a contradiction. □

The results of our lemmas are general, but they have several caveats. For example $S^N(\lambda^N)$ may be empty in Lemma 2 and F^N may be empty in Lemma 3, and the result would still hold. What we are really interested in is to approximate the intersection $S(\lambda) \cap F$, which is not possible if either S^N or F^N are empty with nonzero probability. To preclude that possibility, we effectively need a constraint qualification and perturbation analysis, which will be addressed partly in Section 4 and in future research.

Lemma 4. *Consider the coupled minimization – nonlinear equation problem*

$$\min_{x \in X} f(x,\lambda); \quad g(x,\lambda) = 0.$$

Denote by $S(\lambda)$ the solution set of the first problem and by F the solution set of the second problem. Assume the following

i $\cup_\lambda \{S(\lambda),\lambda\} \cap F$ *has a unique solution point $(x^*,\lambda^*) \in C_X \times C_\Lambda$, where the sets C_X and C_Λ are closed and compact.*

ii *The sequences f^N and g^N converge to f and g uniformly on $C \times C_\Lambda$.*

iii *The coupled problem*

$$\min_{x \in X} f^N(x,\lambda); \quad g^N(x,\lambda) = 0.$$

has a solution $(x^N,\lambda^N) \in C_X \times C_\Lambda$ w.p. 1.

Then, $(x^N,\lambda^N) \to (x^*,\lambda^*)$ w.p. 1.

Proof Define by $S_m^N(\lambda)$ the solution set of the problem, $\min_{x\in X} f^N(x,\lambda)$, and by F^N the solution set of the problem, $g^N(x,\lambda) = 0$. By assumption (iii),

$$x^N \in S_m^N(\lambda^N), (x^N,\lambda^N) \in F^N.$$

Take now an accumulation point of (x^N,λ^N), which we denote by $(\tilde{x},\tilde{\lambda}) \in C_X \times C_\lambda$. Then, by Lemma 2 and using (ii), we have that $S^N(\lambda^N) \to \mathcal{S}(\tilde{\lambda})$ w.p. 1 and, thus, $\tilde{x} \in S(\tilde{\lambda})$. Similarly, by Lemma 3 and using (ii), we have that $(\tilde{x},\tilde{\lambda}) \in F$, and thus

$$\left(\tilde{x},\tilde{\lambda}\right) \in \cup_\lambda \{S(\lambda),\lambda\} \cap F.$$

From assumption (i), we must then have that $\tilde{x} = x^*$ and $\tilde{\lambda} = \lambda^*$. Therefore, (x^N,λ^N) has a unique accumulation point and thus is convergent w.p. 1. □

4. Example application of the theory

We now show how the theory we have developed in Section 3 can be applied to SOESC. The case where Ω is infinite requires too large a technical preamble, so we limit ourselves to the case where the event space is finite, that is, $\Omega = \{\omega_1,\omega_2,\ldots,\omega_M\}$, with $p(\omega_i) = p_i > 0$. In addition, we assume that K_x and K_y represent only nonnegativity constraints. In that case, problem (1) can be rewritten as a deterministic problem:

$$\begin{array}{rl} \min\limits_{x_1\geq 0, y_1(\omega)\geq 0} & \sum_{i=1}^M p_i\phi(x,y(\omega_i),\omega_i) \\ \text{s.t.} & 0 = \sum_{i=1}^M p_i\psi(x,y(\omega_i),\omega_i) \\ \forall \omega_i \in \Omega: & 0 = \Gamma(x,y(\omega_i),\omega_i). \end{array} \tag{6}$$

Here, $y(\omega_i) \in \mathbb{R}^n$, $x \in \mathbb{R}^m$. The mappings ϕ,ψ,Γ are twice continuously differentiable mappings. We denoted by y_1 and by x_1 subsets of the vectors y and x.

To simplify our notation, we use the following aggregate notation:

$$\begin{array}{rcl} Y &=& (y(\omega_1), y(\omega_2),\ldots,y(\omega_M)) \\ Y_1 &=& (y_1(\omega_1), y_1(\omega_2),\ldots,y_1(\omega_M)) \\ z &=& (x,Y), \quad z_1 = (x_1,Y_1) \\ g(z) &=& \begin{pmatrix} \Gamma(x,y(\omega_1),\omega_1) \\ \Gamma(x,y(\omega_2),\omega_2) \\ \vdots \\ \Gamma(x,y(\omega_M),\omega_M) \end{pmatrix}. \end{array}$$

$$\begin{array}{rcl} f(z) &=& \sum\limits_{i=1}^M p_i\phi\left(x,y(\omega_i),\omega_i\right) \\ h(z) &=& \sum\limits_{i=1}^M p_i\psi\left(x,y(\omega_i),\omega_i\right) \end{array}$$

This results in the problem

$$\min_{z,z_1\geq 0} f(z) \text{ s.t } h(z) = 0, g(z) = 0. \tag{7}$$

We assume that its (unique) solution is z^*.

We define by $\mathcal{I}_1$, and, respectively, $\mathcal{I}$, the index set of the constraints $z_1 \geq 0$, which are active at z^* with respect to the vector z_1, and, respectively, z. That is, $z^*_{1,\mathcal{I}_1} = 0$, $z^*_{1,\bar{\mathcal{I}}_1} > 0$, and $z^*_{\mathcal{I}} = 0$. Here, we denote by $\bar{\mathcal{A}}$ the complement of the set $\mathcal{A}$.

Let z^* be the solution of the problem (7). We assume that at z^* we have the following:

[A1] $J(z) = [\nabla_z g \; \nabla_z h]_{\bar{\mathcal{I}}}$ has full column rank at z^*.

[A2] At z^*, we have that

$$u^T \nabla^2_{zz} \left[f + g^T \lambda^* + h^T \mu^* \right] u > 0,$$

$\forall u \neq 0$ that satisfies $\nabla_z g\, u = 0$ and $u_{\mathcal{I}} = 0$.

Here $\lambda^* \in \mathbb{R}^m$, $\mu^* \in \mathbb{R}^p$ are the Lagrange multipliers, which satisfy

$$0 = \nabla_z \left(f + \lambda^{*T} h + \mu^{*T} g \right)_{\bar{\mathcal{I}}} (z^*) \quad (8)$$

and exist from assumption **[A1]**. Assumption **[A1]** is the widely encountered linear independence constraint qualification. Assumption **[A2]**, however, is stronger than the typical second-order condition, in that the positive definiteness of the Lagrangian is required over a larger space, one that excludes the constraints $\nabla_z h\, u = 0$.

Consider now the auxiliary problem

$$\min_z f(z) + h(z)^T \lambda \ \text{ s.t } \ g(z) = 0. \quad (9)$$

We immediately have the following result

Lemma 5. *For $\lambda = \lambda^*$, the solution of problem (9) is z^*. In addition, for $||\lambda - \lambda^*||$ sufficiently small, (9) has a unique solution, $z(\lambda)$, which is continuous in λ. With this notation, λ^* is the only local solution of the nonlinear equation $h(z(\lambda)) = 0$.*

Proof It is immediately recognizable that (8) are precisely the first-order optimality conditions for (9) and that **[A1]** and **[A2]** are the constraint qualification and second-order sufficient conditions for (9). The continuity of $z(\lambda)$ follows from the stability under perturbation of nonlinear programming solutions [4]. From any solution of $h(z(\lambda)) = 0$ we can obtain a point satisfying first-order optimality conditions for (7). But under the conditions here such points are locally unique [4], which completes the proof. $\Box$

Therefore, (z^*, λ^*) is the only local solution of the following mixed minimization – nonlinear equations problem.

$$z = \operatorname*{argmin}_{z, z_1 \geq 0, g(z) = 0} f(z) + h(z)^T \lambda; \quad h(z) = 0. \quad (10)$$

Note that here, in the first minimization λ is a parameter, and not a variable, and $h(z) = 0$ is not enforced.

Lemma 5 then shows the equivalence between the formulation (1) and (2).

The next step is to prove the equivalence of (2) with (4). The key step is the following lemma.

Lemma 6. *Consider the problem*

$$\begin{aligned} \min_{y_1(\omega) \geq 0} \quad & [\phi(x, y(\omega), \omega)] + \lambda^T [\psi(x, y(\omega), \omega)] \\ s.t. \quad & 0 = \Gamma(x, y(\omega), \omega), \end{aligned} \quad (11)$$

for ω one of ω_i, $i = 1, 2, \ldots, M$. Then, in a neighborhood around $x = x^*$ and $\lambda = \lambda^*$, the above problem has a unique solution $y(x, \lambda, \omega)$, which is continuous in (x, λ).

Proof For $\omega = \omega_i$, from (8) it follows that, for $\lambda = \lambda^*$ and $x = x^*$, μ_i^*, the component of μ^* corresponding to $\Gamma(x, y(\omega_i), \omega_i)$ is the Lagrange multiplier of Γ, which satisfies the first-order conditions of (11). Then, from **[A1]** and **[A2]**, it follows that (11) satisfies the constraint qualification and strong second-order sufficient conditions at $y^*(\omega_i)$. From parametric stability results of nonlinear program solutions [4], the conclusion follows. $\Box$.

We are now in position to state and prove our main result.

Theorem 1. *Assume that problem (6) satisfies assumptions* **[A1]** *and* **[A2]** *at $z^* = (x^*, y^*(\omega_1), y^*(\omega_2), \ldots, y^*(\omega_M))$. Assume that SAA applied to (6) results in a solution whose x-value, x^N, and λ-multiplier λ^N (the multiplier of the expectation constraint), exist and are in a compact neighborhood of x^* and λ^*; then x^N converges to x^* with probability one.*

Proof Define by $\tilde{y}(x, \lambda, \omega)$ the solution of problem (11). Define $X = \{x | x_1 \geq 0\}$ and

$$\begin{aligned} \tilde{f}(x, \lambda) &= E\left[\phi(x, y(x, \lambda, \omega), \omega) \right. \\ &\quad \left. + \lambda^T \psi(x, y(x, \lambda, \omega), \omega)\right] \\ \tilde{g}(x, \lambda) &= E\left[\psi(x, y(x, \lambda, \omega), \omega)\right]. \end{aligned}$$

From Lemma 5 (with $h(z(\lambda)) = \tilde{g}(x, \lambda)$), it follows that the solution of (6) is the unique local solution of the mixed optimization – nonlinear equation problem

$$\min_{x \in X} \tilde{f}(x, \lambda); \quad \tilde{g}(x, \lambda) = 0.$$

It immediately follows that assumption (i) of Lemma 4 is satisfied. If i.i.d. sampling is used, assumption (ii) of that lemma follows as well [7]. Finally, we note than any first–order point of the SAA problem attached to (6) is also a solution of the SAA mixed optimization–nonlinear equation problem, thus assumption (iii) of Lemma 4 is satisfied. Lemma 4 can be applied to obtain the conclusion. □

Perhaps the most arguable assumptions of our approach are assumptions [**A2**] and (iii) in Lemma 4. We point out that [**A2**] is satisfied in the case where ϕ is a strongly convex function in x and y uniformly in ω and Γ and ψ are linear in x and y. We envision that the assumption (iii) can be relaxed in that the SAA mixed optimization–nonlinear equation problem does not need to have a local solution with probability one, but only with a probability that asymptotically converges to one.

5. Conclusions and future work

In this work we have presented a convergence framework for SAA approaches for SOESC. Such problems have recently been formulated in a variety of applications.

In future work, optimality conditions of generic SOESC will be studied. This will allow us to extend the SAA convergence result to problems with infinite Ω event spaces. It seems reasonable from our analysis that all that is needed is the existence of a Lagrangian with slightly stronger second-order conditions (i.e., its Hessian is positive definite on a slightly larger space than for the sufficient conditions). Nonetheless, the technical aspects are far beyond a short communication like this one.

Another interesting extension is to consider a saddle point problem instead of an optimization problem in Lemma 4, which could remove the assumption of the boundedness below of the optimization problem (3). In turn, this would avoid the need for strong second-order conditions.

References

[1] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*, Springer, New York, 1997.

[2] G. Gurkan, A. Ozge, and S. M. Robinson, *Sample-path solution of stochastic variational inequalities*, Mathematical Programming, 84 (1999), pp. 313–333.

[3] P. Krokhmal, J. Palmquist, and S. Uryasev, *Portfolio optimization with conditional value-at-risk objective and constraints*, JOURNAL OF RISK, 4 (2002), pp. 43–68.

[4] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Series in Operations Research, Springer, 1999.

[5] E. Plambeck, B. Fu, S. Robinson, and R. Suri, *Sample-path optimization of convex stochastic performance functions*, Mathematical Programming, 75 (1996), pp. 137–176.

[6] J. Primbs, *Stochastic receding horizon control of constrained linear systems with state and control multiplicative noise*, in American Control Conference, 2007. ACC'07, 2007, pp. 4470–4475.

[7] A. Shapiro and A. Ruszczynski, *Lectures on stochastic programming*, 2007. Book in progress. Available at www2. isye. gatech. edu/ ashapiro/publications. html, Nov.

[8] W. Wang and S. Ahmed, *Sample average approximation of expected value constrained stochastic programs*, Operations Research Letters, 36 (2008), pp. 515 – 519.