An Interior Decomposition Algorithm for Two-Stage Stochastic Convex Program, and Integration Formulae and Scenario Generation via Optimization

Sanjay Mehrotra

Northwestern University,
Department of Industrial Engineering and Management Sciences

DOE Applied Mathematics Meeting, October 17–19, 2011
Award: DE-FG02-10ER26037; SC0005102.
Title: "Sparse Grid Scenario Generation and Interior Algorithms for Stochastic Optimization in a Parallel Computing Environment"

Key accomplishments to date

1. Development of a new interior decomposition algorithm
   - suitable for a multi-core, massively parallel environment
   - where some computational nodes might fail
2. Theoretical and empirical justification of the use of sparse grid methods in scenario generation
3. Development of a new algorithm to generate scenarios with nonnegative weights, matching any prescribed set of moments
   - nonnegative weights have theoretical advantage when used in a convex optimization framework
   - potentially useful for problems with many random variables with structured uncertainty

The convex stochastic two-stage problem

- Convex stochastic two-stage optimization problem with K scenarios

$$\min \; c^T x + \sum_{k=1}^{K} \bar{\eta}^k(x) \text{ s.t. } x \in G \cap L$$

$$\bar{\eta}^k(x) := \min \; (d^k)^T y^k \text{ s.t. } y^k \in G^k \cap L^k(x)$$

- Objectives:
  - Self-concordance of barrier formulations?
  - Find a general decomposition method

The two-stage barrier problem

- Let b and $B^k, k = 1, \ldots, K$ be non-degenerate and strongly self-concordant barrier functions on int G and int G^k, resp.
- The two-stage barrier problem is defined as

$$\min\ f(x,\mu) := c^T x + \mu b(x) + \sum_{k=1}^{K} \eta^k(x,\mu) \ \text{ s.t. } x \in L, \quad \text{(TSBP)}$$

$$\eta^k(x,\mu) := \min\ {d^k}^T y^k + \mu B^k(y^k) \ \text{ s.t. } y^k \in L^k(x), \quad \text{(SSBP)}$$

Theorem

The barrier recourse functions $\eta^k(x,\mu)$ are differentiable in x and μ; convex in x and concave in μ; self-concordant in x. The family $\{\eta^k(x,\mu), \mu > 0\}$ is self-concordant.

Theorem

The composite barrier function $f(x,\mu)$ is self-concordant; and $\{f(x,\mu), \mu > 0\}$ is a self-concordant family.

Interior Point Decomposition Algorithms

- Derived a short-step and a long-step interior point algorithm from the barrier formulation
- Convergence of both were proven
  - short-step iteration complexity is $O(\sqrt{\tilde{\vartheta}} \ln \mu^0/\epsilon)$,
  - long-step iteration complexity is $O(\tilde{\vartheta} \ln \mu^0/\epsilon)$,
  - where $\tilde{\vartheta} = \sqrt{\frac{1}{\mu}\tilde{\Delta}x^T \nabla_x^2 f(x,\mu)\tilde{\Delta}x}$;
- This general result matches the known iteration complexity of several special cases
  - two-stage linear stochastic programming
  - two-stage quadratic stochastic programming
  - two-stage semidefinite stochastic programming

The scenario generation problem I.

- A general stochastic program:

$$\min_{x \in \mathcal{X}} \int_{\Xi} f(x, \xi) \mu(d\xi),$$

- Two-stage problems: $f(\cdot, \xi)$ is the optimal value function of the second stage; evaluating it is expensive
- Scenario generation:

$$\int_{\Xi} f(x, \xi) \mu(d\xi) \approx \sum_{k=1}^{K} w_k f(x, \xi_k)$$

- Scenario generation $\equiv$ cubature formulas of numerical integration

The scenario generation problem II.

$$\int_{\Xi} f(x,\xi)\mu(d\xi) \approx \sum_{k=1}^{K} w_k f(x,\xi_k)$$

Desirable properties of formulas:

- Good approximation
  - "Moment matching": the formula is exact for every polynomial f up to a certain degree
- Small K (number of scenarios)
  - Fewer than usual in numerical integration
- Nonnegative weights
  - Yields convex approximations of convex left-hand sides
  - If $f \geq 0$, evaluating the right-hand side is not prone to cancellations
- Prescribed domain

Moment matching

- Goal: make the approximation exact for a set of polynomials
  - by extension, exact for all linear combinations
  - notation: $u_x = (p_1(x), \ldots, p_N(x))$
- Example 1: all monomials up to a certain degree,

$$u_x = (1, x_1, x_2, \ldots, x_n^d)$$

- Example 2: all monomials up to degree d, and all univariate polynomials up to degree D.
- Moment matching formula:

$$\sum_{k=1}^{K} w_k u_{\xi_k} = m,$$

where m is the vector of integrals of the components of u_x.

Moment matching and column generation

- Moment matching is a semi-infinite LP (feasibility problem).
- Given nodes $\xi_1, \dots, \xi_\ell$ finding the weights $w_1, \dots, w_\ell$ is an LP
- Weighing the nodes:

$$\min_{w \in \mathbb{R}^\ell,\, \alpha \in \mathbb{R}^N} \left\{ \sum_{i=1}^{N} |\alpha_i| \;\middle|\; \sum_{k=1}^{\ell} w_k u_{\xi_k} + \alpha_i = m,\ w \geq 0 \right\}$$

Lemma

If optimal value is 0, a solution is found. Otherwise the reduced cost of "column" u_ξ is $-p^*(\xi)$, where p^* is the polynomial whose coefficient vector (in the u_x basis) is the dual optimal vector.

- Finding the column with the most negative reduced cost amounts to a polynomial optimization problem.

Column generation oracles

Theorem

Suppose we are given an oracle that finds a node $\xi_{\ell+1}$ with strictly negative reduced cost, given the nodes $\{\xi_1, \ldots, \xi_\ell\}$ and the optimal solution to the corresponding LP above. Using this oracle, a positive formula can be found in oracle-polynomial time.

- Global polynomial optimization is NP-hard even for low degree
- To find a point ξ for which $p^*(\xi) > 0$ can be done with random sampling:

Lemma

Let p^* be the optimal dual solution with objective function value $I > 0$. Let B be an upper bound on the maximum of the dual feasible polynomials. Let $x \leq \min(B, I)$, and draw random points from Ξ with the distribution determined by μ. Then the expected number of points ξ needed to be drawn until one that satisfies $p^*(\xi) \geq x$ is found is at most $(B - x)/(I - x)$.

Computational results I – integrals

Approximation error for difficult integrals

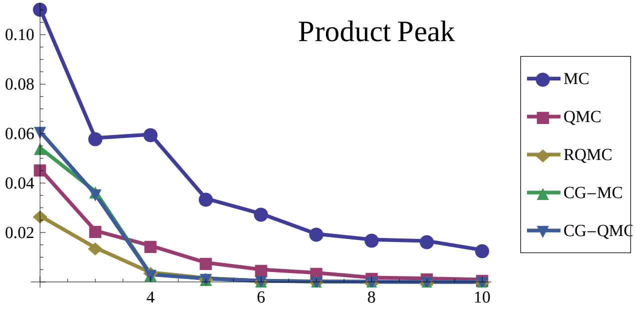

Figure: Performance profiles of five cubature formulas: Monte Carlo (MC), quasi-Monte Carlo (QMC), reweighed QMC (RQMC), and column generation using MC and QMC sampling (CG-MC, CG-QMC). Horizontal axis: degree of exactness of CG-MC, CG-QMC, and RQMC methods. Vertical axis: median relative errors from 200 experiments with a four-variate parametric family.

Computational results II – optimization

Utility maximization model from [Pennanen-Koivu]:

$$\min_{x \in \mathcal{X}} \int_{\mathbb{R}^n} \exp(-\xi^{\mathrm{T}} x)\mu(d\xi), \quad \mathcal{X} = \left\{ x \in \mathbb{R}^n_+ \,\middle|\, \sum_i x_i \leq 1 \right\},$$

d	K	MC	QMC	CG-MC	CG-QMC	RQMC
2	28	0.1994	0.1817	0.0683	0.0102	0.0597
3	84	0.1139	0.1130	0.0037	0.0726	0.0486
4	210	0.0661	0.0626	0.0057	0.0015	0.0187
5	462	0.0457	0.0319	0.0010	0.0019	0.0136
6	924	0.0299	0.0189	0.0070	0.0028	0.0001
7	1716	0.0245	0.0078	0.0044	0.0037	0.0030

Table: Relative errors of the approximate solutions to a utility maximization model, as a function of the degree of exactness d and the number of scenarios $K = \binom{d+6}{6}$. Acronyms are on previous slide.

Publications

Papers acknowledging the grant are in the pipeline:

- Chen, M., Mehrotra, S.
  Self-concordance and Decomposition Based Interior Point Methods for Stochastic Convex Optimization Problem
  To appear in SIAM Journal on Optimization.
- Chen, M., Mehrotra, S.
  Scenario Generation for Stochastic Problems via the Sparse Grid Method
  Technical report (under revision).
- Mehrotra, S., Papp, D.
  Generating Moment-matching Scenarios Using Optimization Techniques
  In preparation.