Errata to "A scenario decomposition algorithm for 0-1 stochastic programs," [Oper. Res. Lett. 41(6):565-569, 2013]

Shabbir Ahmed
School of Industrial & Systems Engineering,
Georgia Institute of Technology, 765 Ferst Drive, Atlanta, GA 30332.

November 25, 2014

Section 3 of the paper "A scenario decomposition algorithm for 0-1 stochastic programs," *Operations Research Letters*, 41(6):565-569, 2013, contains some errors. The result in Proposition 2 had to be stated with respect to epsilon optimality as opposed to optimality. In this errata the entire Section 3 is reproduced with the errors corrected. The portions highlighted in red indicate the changes with respect to the original paper.

3. Optimality of scenario solutions

The algorithm proposed in the previous section explores solutions to the scenario subproblems as candidates for solutions to the overall problem. This is a departure from other methods based on solving the dual problem (3) where candidate solutions are generated by aggregating solutions to the scenario subproblems. In this section we attempt to provide some rationale why the solutions to the scenario subproblems themselves can be good solutions to the overall problem in the case the problem of interest is a sample average approximation problem.

Consider a 0-1 stochastic program

$$(P): \ \min \{\mathbb{E}[f(x, \boldsymbol{\xi})] : \ x \in X \subseteq \{0,1\}^n\},$$

and its sample average approximation

$$(SAA_N): \ \min \left\{(1/N)\sum_{i=1}^{N} f(x, \xi^i) : \ x \in X \subseteq \{0,1\}^n\right\},$$

corresponding to an iid sample $\{\xi^i\}_{i=1}^N$. The scenario subproblem corresponding to the i-th scenario/sample is

$$(P_i): \ \min \left\{f(x, \xi^i) : \ x \in X \subseteq \{0,1\}^n\right\}.$$

Let S_N^ϵ be the set of ϵ-optimal solutions of (SAA_N) for a given $\epsilon > 0$, i.e. $S_N^\epsilon = \{x \in X : (1/N)\sum_{i=1}^N f(x, \xi^i) \leq (1/N)\sum_{i=1}^N f(y, \xi^i) + \epsilon \ \ \forall y \in X\}$. Let S_i be the set of optimal solutions of (P_i) for $i = 1, \ldots, N$. Note that these are random sets since they depend on the sample drawn. We investigate the following question:

> What is the probability that the set of solutions to one of the scenario problems (P_i) contains an ϵ-optimal solution to (SAA_N)?

More precisely we would like to estimate $\Pr[S_N^\epsilon \cap (\bigcup_{i=1}^N S_i) \neq \emptyset]$ and understand its dependence on X and N.

Trivially, if $|X| \leq 2$ and $N \geq 2$ then the above probability is 1. On the other hand, it is not hard to construct examples where this probability is arbitrarily small. For example, suppose $X = \{x^1, x^2, x^3\}$, $\boldsymbol{\xi}$ has two realizations ξ^1 and ξ^2 with equal probability, and the values of $f(x, \xi)$ are as in Table 1. Then for

Solution	ξ^1	ξ^2
x^1	0	0
x^2	-1	2
x^3	2	-1

Table 1: Values of $f(x, \xi)$ for an example

N reasonably large and ϵ sufficiently small, $S_N^\epsilon = \{x^1\}$ with very high probability while $\bigcup_{i=1}^N S_i = \{x^2, x^3\}$ with probability one.

In the following we show that if the collection of random objective functions $\{f(x, \boldsymbol{\xi})\}_{x \in X}$ of (P) is jointly normal with a nonnegative and strictly diagonally dominant covariance matrix then

$$\lim_{N\to\infty} \Pr\left[S_N^\epsilon \cap (\bigcup_{i=1}^N S_i) \neq \emptyset\right] = 1$$

exponentially fast. Before proceeding with a proof of the above result we make some remarks. The condition of nonnegative covariance matrix implies that the objective functions at different solutions are nonnegatively correlated, and therefore, perhaps it is not very surprising that a scenario solution is likely to be near optimal for the sample average problem. Recall that a matrix is strictly diagonally dominant if the diagonal element in a row is greater than the sum of the absolute values of all other elements in that row. This condition can be ensured by adding independent zero-mean normal noise to the objective function value for each solution without changing the optimal solution or optimal value of (P). We will need the following Lemma.

Lemma 1 *Let W be a $K+1$ dimensional Gaussian random vector, $W \sim N(\mu, \Sigma)$ with the following properties:*

1. *For some $\epsilon \geq 0$, $\mu_0 \leq \mu_k + \epsilon$ for all $k = 1, \ldots, K$ and*
2. *the covariance matrix Σ is nonnegative and strictly diagonally dominant.*

Condition 2 implies that there exists δ such that $0 < \delta^2 := \min_{k \neq l}\{\sigma_k^2 - \sigma_{kl}, \sigma_l^2 - \sigma_{kl}\}$. Then

$$\Pr[W_0 \leq W_k \; \forall k = 1, \ldots, K] \geq (\Phi(-\sqrt{2}\epsilon/\delta))^K,$$

where Φ is the standard normal cdf.

Proof: Since $\mu_0 \leq \mu_k + \epsilon$ for all $k = 1, \ldots, K$, we have

$$\Pr[W_0 \leq W_k \; \forall k = 1, \ldots, K] \quad \geq \quad \Pr[(W_0 - \mu_0) - (W_k - \mu_k) \leq -\epsilon \; \forall k = 1, \ldots, K].$$

Let $U_k := (W_0 - \mu_0) - (W_k - \mu_k)$ for $k = 1, \ldots, K$, and note that U is a K dimensional Gaussian random vector with $\mathbb{E}[U_k] = 0$, $\mathbb{E}[U_k^2] = \sigma_0^2 + \sigma_k^2 - 2\sigma_{0k} \geq 2\delta^2$, and $\mathbb{E}[U_k U_l] = \sigma_0^2 - \sigma_{0l} - \sigma_{0k} + \sigma_{kl} \geq 0$ for all $k, l = 1, \ldots, K$. Consider another K dimensional Gaussian random vector V with $\mathbb{E}[V_k] = 0$, $\mathbb{E}[V_k^2] = \mathbb{E}[U_k^2]$ and $\mathbb{E}[V_k V_l] = 0$ for all $k, l = 1, \ldots, K$. Since $\mathbb{E}[V_k V_l] \leq \mathbb{E}[U_k U_l]$ for all $k, l = 1, \ldots, K$, by Slepian's inequality [14]

$$\Pr[W_0 \leq W_k \; \forall k] \geq \Pr[U_k \leq -\epsilon \; \forall k] \geq \Pr[V_k \leq -\epsilon \; \forall k] = \prod_{k=1}^K \Pr[V_k \leq -\epsilon],$$

where the last identity follows from independence. Since $\mathbb{E}[V_k^2] \geq 2\delta^2$. Hence for any k,

$$\Pr[V_k \leq -\epsilon] = \Pr\left[V_k/\sqrt{\mathbb{E}[V_k^2]} \leq -\epsilon/\sqrt{\mathbb{E}[V_k^2]}\right] \geq \Pr\left[V_k/\sqrt{\mathbb{E}[V_k^2]} \leq -\sqrt{2}\epsilon/\delta\right] = \Phi(-\sqrt{2}\epsilon/\delta)$$

and the result follows. □

Proposition 2 *If the collection of random objective functions $\{f(x,\boldsymbol{\xi})\}_{x\in X}$ of (P) is jointly normal with a nonnegative and strictly diagonally dominant covariance matrix then for any $\epsilon > 0$,*

$$\lim_{N\to\infty} \Pr\left[S_N^\epsilon \cap (\bigcup_{i=1}^N S_i) \neq \emptyset\right] = 1$$

exponentially fast.

Proof: Let $\alpha = \epsilon/2$ and S_*^α be the set of α-optimal solutions to (P). Note that,

$$\Pr\left[S_N^\epsilon \cap (\bigcup_{i=1}^N S_i) \neq \emptyset\right] \geq \Pr\left[[S_*^\alpha \cap (\bigcup_{i=1}^N S_i) \neq \emptyset] \wedge [S_*^\alpha \subseteq S_N^\epsilon]\right] \geq \Pr[S_*^\alpha \cap (\bigcup_{i=1}^N S_i) \neq \emptyset] + \Pr[S_*^\alpha \subseteq S_N^\epsilon] - 1,$$

where the last inequality is using Boole's inequality.

Let $f_N(x) = (1/N)\sum_{i=1}^N f(x,\xi^i)$ and $f(x) = \mathbb{E}[f(x,\boldsymbol{\xi})]$. Note that

$$\begin{aligned} S_*^\alpha \not\subseteq S_N^\epsilon &\Rightarrow \exists\, x,y \in X \text{ s.t. } f(x) \leq f(y) + \alpha,\ f_N(x) > f_N(y) + \epsilon \\ &\Rightarrow \exists\, x,y \in X \text{ s.t. } f_N(x) - f(x) > f_N(y) - f(y) + \epsilon/2 \\ &\Rightarrow \exists\, x,y \in X \text{ s.t. } g_N(x,y) - g(x,y) > \epsilon/2, \end{aligned}$$

where $g_N(x,y) = f_N(x) - f_N(y)$ and $g(x,y) = f(x) - f(y)$. Note that $\mathbb{E}[g_N(x,y)] = g(x,y)$. Let $g(x,y,\boldsymbol{\xi}) = f(x,\boldsymbol{\xi}) - f(y,\boldsymbol{\xi})$. Applying Large deviations theory (cf. [8]) to $g(x,y,\boldsymbol{\xi})$ we have that

$$\Pr\left[g_N(x,y) - \mathbb{E}[g_N(x,y)] > \epsilon/2\right] \leq e^{-N\frac{\epsilon^2}{8\sigma_{\max}^2}}$$

where $\sigma_{\max}^2 = \max_{x,y\in X}[\mathbb{V}[g(x,y,\boldsymbol{\xi})]$. Thus

$$\Pr[S_*^\alpha \not\subseteq S_N^\epsilon] \leq \sum_{x\in X}\sum_{y\in X} \Pr\left[g_N(x,y) - \mathbb{E}[g_N(x,y)] > \epsilon/2\right] \leq |X|^2 e^{-N\frac{\epsilon^2}{8\sigma_{\max}^2}}.$$

Thus

$$\Pr[S_*^\alpha \subseteq S_N^\epsilon] \geq 1 - |X|^2 e^{-N\frac{\epsilon^2}{8\sigma_{\max}^2}}.$$

Now consider $x^* \in S_*^\alpha$. We have

$$\begin{aligned} \Pr[S_*^\alpha \cap (\textstyle\bigcup_{i=1}^N S_i) \neq \emptyset] &\geq \Pr[x^* \in (\textstyle\bigcup_{i=1}^N S_i)] \\ &= 1 - \textstyle\prod_{i=1}^N \Pr[x^* \notin S_i] \\ &= 1 - \textstyle\prod_{i=1}^N (1 - \Pr[x^* \in S_i]) \\ &= 1 - \textstyle\prod_{i=1}^N (1 - \Pr[f(x^*,\boldsymbol{\xi}) \leq f(y,\boldsymbol{\xi})\ \forall\, y \in X \setminus \{x^*\}]). \end{aligned}$$

Applying Lemma 1 with $K = |X \setminus \{x^*\}|$,$W_0 = f(x^*,\boldsymbol{\xi})$ and $W_k = f(y,\boldsymbol{\xi})$ for $y \in X \setminus \{x^*\}$, we get

$$\Pr\left[f(x^*,\boldsymbol{\xi}) \leq f(y,\boldsymbol{\xi})\ \forall\, y \in X \setminus \{x^*\}\right] \geq (\Phi(-\sqrt{2}\alpha/\delta))^{|X|},$$

for some $\delta > 0$ depending on the covariance matrix of $\{f(x,\boldsymbol{\xi})\}_{x\in X}$. Thus

$$\Pr\left[S_N^\epsilon \cap (\bigcup_{i=1}^N S_i) \neq \emptyset\right] \geq (1 - (1 - (\Phi(-\sqrt{2}\epsilon/2\delta))^{|X|})^N) - |X|^2 e^{-N\frac{\epsilon^2}{8\sigma_{\max}^2}}. \tag{1}$$

Since $0 < \Phi(-\sqrt{2}\epsilon/2\delta) < 1$ for $\delta > 0$, taking limits with respect to N on both sides of (1) we get the desired result. $\square$

We close this section with some remarks on Proposition 2. Note that when ϵ is very small relative to δ, $\Phi(-\sqrt{2}\epsilon/2\delta) \approx 1/2$ and so the right hand side of (1) is approximately $(1-(1-(1/2)^{|X|})^N) - |X|^2 e^{-NC}$ where

C is a small constant. This indicates that smaller the solution set,$|X|$, relative to the sample size N, higher is the probability that one of the scenario solutions is optimal. If $|X|$ is large relative to N, then the bound in (1) is very loose, even negative. Consequently inequality (1) does not lead to any meaningful finite sample size estimates, but implies asymptotic behavior with respect to N for fixed X. Proposition 2 can be extended to more general distributions under different assumptions. For example, if the random vector $\{f(x^*, \boldsymbol{\xi}) - f(y, \boldsymbol{\xi})\}_{y \in X \setminus \{x^*\}}$ (where x^* is an ϵ-optimal solution of (P)) has a density and contains a ball centered at 0 of radius at least ϵ in its support, then there exists $p > 0$ such that $\Pr\left[f(x^*, \boldsymbol{\xi}) \leq f(y, \boldsymbol{\xi}) \; \forall \; y \in X \setminus \{x^*\}\right] \geq p$. We can then replace $(\Phi(-\sqrt{2}\epsilon/2\delta))^{|X|}$ with p in (1) and the conclusion of Proposition 2 holds.