Model Predictive Control and Stochastic Dynamic Programming

Dominic Keehan[1], Andy Philpott[1], Eddie Anderson[2]

[1]Electric Power Optimization Centre,
University of Auckland

[2]Imperial College London and University of Sydney

XVI ICSP, UC Davis, CA
July 2023

Intro

When one has access to samples of the random variables involved, multistage stochastic optimization problems can be approached by:

- **S** Solving a problem formulated in a scenario tree based on sample average approximation.
- **M** Solving a problem where the random variables in each stage are fixed at their sample mean.

Stochastic programmers consider approach **M** to be a cardinal sin. In this talk we show that its use can sometimes be forgiven.

An example

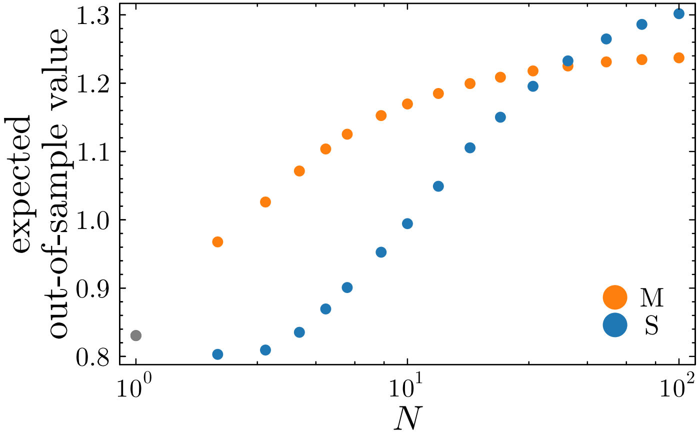
Consider a vendor with an inventory of stock that wishes to maximize their β-discounted expected reward, where they

0. Create policies **S** and **M** using N sample prices.
1. Observe a random price from the distribution $\mathbb{P}$.
2. Using **S** or **M** choose an amount of stock to sell at this price.
3. Pay a storage cost on their remaining inventory.

   Repeat 1–3...

We compare the performance of **S** and **M** in terms of their expected out-of-sample value evaluated with $\mathbb{P}$. When only one sample is available, **S** and **M** are the same policy. One might expect **S** to outperform **M** as N increases.

This is not always the case...



$\beta = 0.95$, $x_0 = 1$, $C(x) = \frac{1}{2}x^2$, $\mathbb{P} = \text{LogNormal}\left(\mu = -\frac{1}{2}, \sigma^2 = 1\right)$

Background

S is an approach that is based on stochastic dynamic programming, and **M** is an approach that is based on model predictive control.

Stochastic dynamic programming does not scale well. Model predictive control is used as a simplified practical alternative.

And yet, some computational studies (e.g. Martin, 2021) show model predictive control performing well out-of-sample.

A stochastic inventory control problem

The inventory control problem introduced earlier is formally

$$\text{SIC:} \quad \max \mathbb{E}\left[\sum_{t=1}^{\infty} \beta^{t-1} \left(P_t u_t - C(x_t)\right)\right]$$

where at each stage t stock u_t is sold from the current inventory x_{t-1} at the observed market price p_t and storage cost $C(x_{t-1} - u_t)$ is paid on the remaining inventory. Each x_t and u_t satisfy

$$\begin{aligned} x_t &= x_{t-1} - u_t, \quad t = 1, 2, \ldots \\ u_t &\in [0, x_{t-1}], \quad t = 1, 2, \ldots, \end{aligned}$$

u_t is nonanticipative, and $x_0 \geq 0$.

Simplifying assumptions

To analyze SIC, we make the following assumptions:

A1 The discount factor $\beta \in (0, 1)$.

A2 The prices P_t are IID, with $P_t \sim \mathbb{P}$ and $\mathrm{supp}(\mathbb{P})$ bounded.

A3 The inventory cost $C : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is an increasing strictly convex and continuously differentiable function with $C(0) = 0$ and derivative c such that $\lim_{x\to\infty} c(x) = \infty$.

Closed-form optimal policy

Theorem (9.2, Stokey, Lucas, and Prescott, 1989)

Let

$$V(x, p) := \max_{0 \leq u \leq x} \{pu - C(x - u) + \beta \mathbb{E}[V(x - u, P)]\}.$$

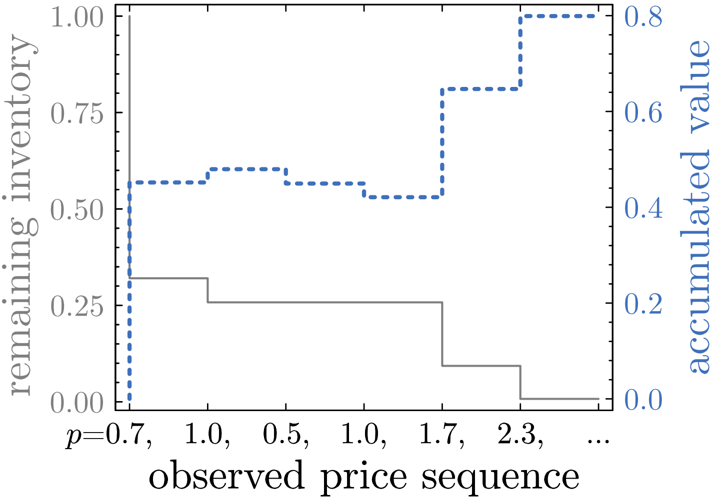
Under A1–A3, SIC has optimal value $\mathbb{E}[V(x_0, P)]$.

Proposition

Under A1–A3 for inventory x and observed price p the optimal policy for problem SIC is to sell

$$u(x, p) = x - c^{-1}\left((\beta(\mathbb{E}[(P - p)_+] + p) - p)_{[c(0), c(x)]}\right).$$

An example trajectory



Sample-based policies

For sample prices $q_1, q_2, \ldots, q_N$ the policy **S** sells

$$u_{\mathbf{S}}(x,p) = x - c^{-1}\left(\left(\beta\left(\frac{1}{N}\sum_{i=1}^{N}(q_i - p)_+ + p\right) - p\right)_{[c(0),c(x)]}\right)$$

and the policy **M** sells

$$u_{\mathbf{M}}(x,p) = x - c^{-1}\left(\left(\beta\left(\left(\sum_{i=1}^{N}\frac{1}{N}q_i - p\right)_+ + p\right) - p\right)_{[c(0),c(x)]}\right).$$

Observe that

$$\frac{1}{N}\sum_{i=1}^{N}(q_i - p)_+ \geq \left(\sum_{i=1}^{N}\frac{1}{N}q_i - p\right)_+.$$

It follows that $u_{\mathbf{M}}(x,p) \geq u_{\mathbf{S}}(x,p)$, and the current inventory of the vendor using **M** is always less than or equal to that using **S**.

Out-of-sample performance

We compare the performance of **S** and **M** analytically by solving:

$$\bar{V}_{\mathbf{S}}(x) = \mathbb{E}\left[Pu_{\mathbf{S}}(x,P) - C(x - u_{\mathbf{S}}(x,P)) + \beta\bar{V}_{\mathbf{S}}(x - u_{\mathbf{S}}(x,P))\right]$$

and

$$\bar{V}_{\mathbf{M}}(x) = \mathbb{E}\left[Pu_{\mathbf{M}}(x,P) - C(x - u_{\mathbf{M}}(x,P)) + \beta\bar{V}_{\mathbf{M}}(x - u_{\mathbf{M}}(x,P))\right].$$

When looking at out-of-sample values, we replace A2 with:

A2' $\mathbb{P}$ has: $\mathrm{supp}(\mathbb{P}) \subseteq \mathbb{R}_+$, $\mathbb{E}_{\mathbb{P}}[P]$ finite, and no atoms.

Under assumptions A1–A3 the values $\bar{V}_{\mathbf{S}}(x_0)$ and $\bar{V}_{\mathbf{M}}(x_0)$ are the out-of-sample values of the policies **S** and **M** respectively.

An outlier-based criterion

Definition

Let $p_{\mathbf{S}}(x)$ be the p that solves

$$\beta(\tfrac{1}{N}\textstyle\sum_{i=1}^{N}(q_i - p)_+ + p) - p = c(x).$$

$p_{\mathbf{S}}(x)$ is the threshold price for inventory x below which $u_{\mathbf{S}}(p) = 0$.

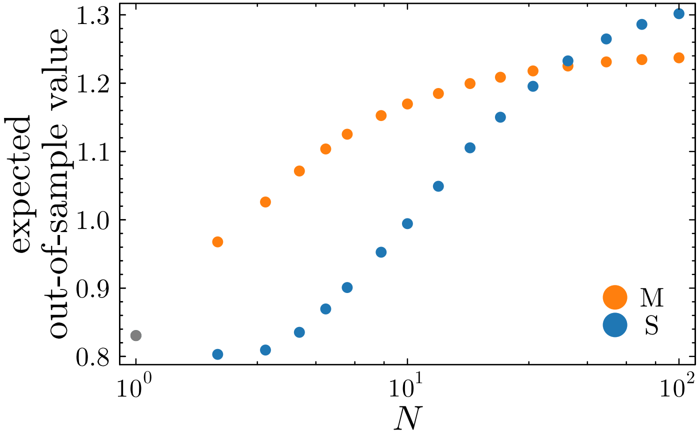
Proposition

Assume that $\mathbb{P}$ has a density f. Under assumptions A1–A3, if

$$c(x) \geq \beta \int_{p_{\mathbf{S}}(x)}^{\infty} p\, f(p) dp$$

for all $x \in [0, x_0]$, then $\bar{V}_{\mathbf{M}}(x) \geq \bar{V}_{\mathbf{S}}(x)$.

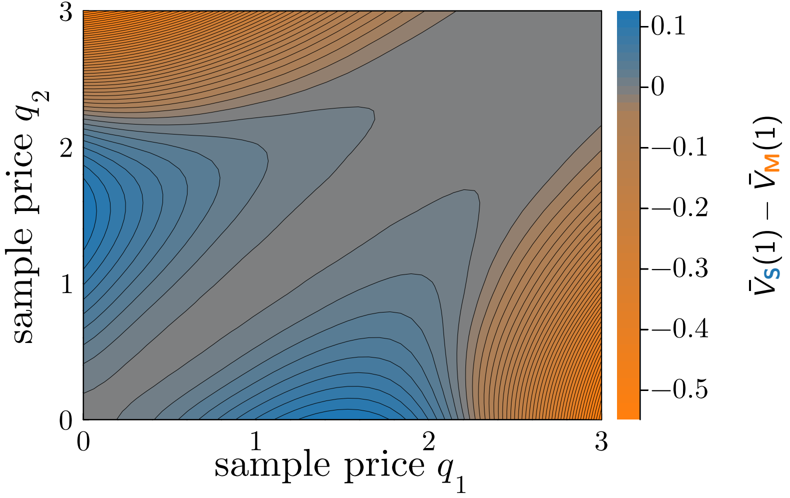
Simulation with a right-skewed price distribution



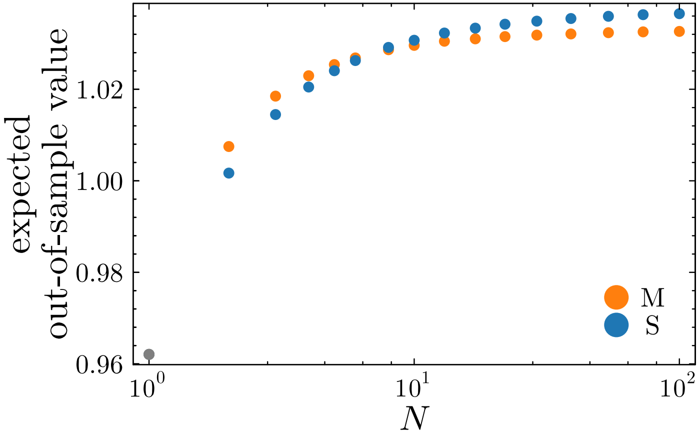
$\beta = 0.95,\ x_0 = 1,\ C(x) = \frac{1}{2}x^2,\ P \sim \text{LogNormal}\left(\mu = -\frac{1}{2}, \sigma^2 = 1\right).$

Comparison of S and M as a function of the samples

Fixing $N = 2$ and varying the sample prices q_1 and q_2

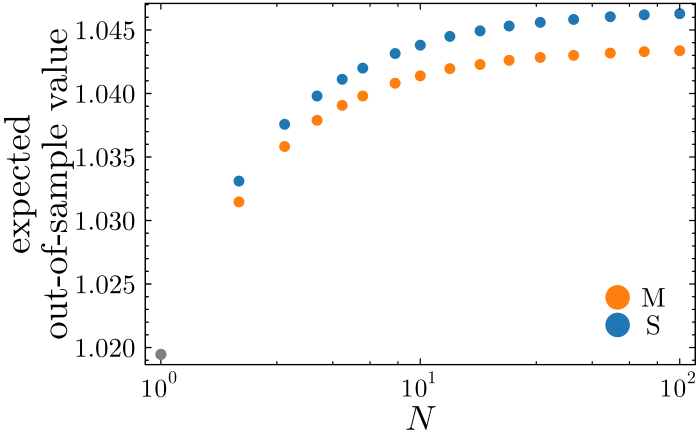


Simulation with another right-skewed price distribution



$P \sim$ Triangular $\left(\frac{1}{2}, 2, \frac{1}{2}\right)$.

Simulation with a left-skewed price distribution



$P \sim$ Triangular $\left(0, \frac{3}{2}, \frac{3}{2}\right)$.

Limiting behaviour of sample average approximation

Right-skewed distributions make **M** better than **S** for small N. In our previous examples this eventually disappears for large N.

Proposition

If $\mathbb{P} = \text{Exponential}(\lambda = 1)$, $N \geq 2$, *and* $C(x) = \frac{1}{2}x^2$, *then as* $\beta \to 1$, *the difference* $\mathbb{E}[\bar{V}_{\mathbf{S}}(1) - \bar{V}_{\mathbf{M}}(1)] \to -\infty$, *where the expectation is over the samples* $q_1, q_2, \ldots, q_N$.

Without $\text{supp}(\mathbb{P})$ bounded, we can make this N arbitrarily large by choosing a β sufficiently close to 1. When $\beta = 1$, the expected out-of-sample value $\mathbb{E}[\bar{V}_{\mathbf{S}}(1)]$ is not consistent.

A distributionally robust interpretation

Let $\mathcal{P}(\mathbb{R})$ denote the set of probability distributions with support on $\mathbb{R}$. Define $\mathcal{M}_1(\mathbb{P}) := \{\mathbb{Q} \in \mathcal{P}(\mathbb{R}) : \mathbb{E}_{\mathbb{Q}}[P] = \mathbb{E}_{\mathbb{P}}[P]\}$. Then

Proposition

The distributionally robust functional equation

$$V_R(x,p) = \max_{0 \le u \le x} \left\{ pu - C(x-u) + \min_{\mathbb{Q} \in \mathcal{M}_1(\mathbb{P})} \beta \mathbb{E}_{\mathbb{Q}}[V_R(x-u, P)] \right\}$$

is satisfied by the solution to

$$V_M(x,p) = \max_{0 \le u \le x} \{ pu - C(x-u) + \beta V_M(x-u, \mathbb{E}_{\mathbb{P}}[P]) \}.$$

References

Dominic Keehan, Andy Philpott, and Edward Anderson (May 2023). "Sample average approximation and model predictive control for inventory optimization". Preprint.

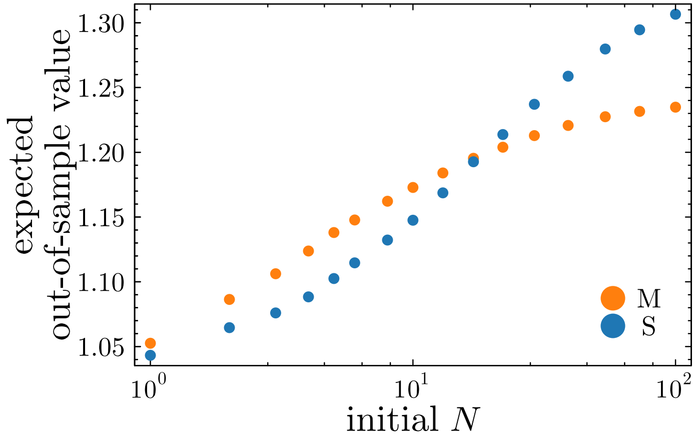
Nancy L. Stokey, Robert E. Lucas, and Edward C. Prescott (1989). "Stochastic Dynamic Programming". In: **Recursive Methods in Economic Dynamics**. Harvard University Press, pp. 239–287.

Thomas Martin (2021). "Stochastic optimization for the procurement of crude oil in refineries". PhD thesis. École des Ponts ParisTech.

Extra slide

Rolling horizon simulation

In practice, one would add each observed sample price to **S** and **M**



$\beta = 0.95,\ x_0 = 1,\ C(x) = \frac{1}{2}x^2,\ P \sim \text{LogNormal}\left(-\frac{1}{2}, 1\right).$