SDDP and Model Predictive Control

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

1. Electric Power Optimization Centre, University of Auckland
2. University of Sydney/Imperial College

October 17, 2022

SDDP and MPC

- SDDP (Stochastic Dual Dynamic Programming)
  - Sample random variables to yield finite scenario tree (SAA).
  - Approximately solve SAA problem using cutting planes.
- MPC (Model Predictive Control)
  - Estimate expectations of random variables and solve deterministic DP.
  - Implement stage 1 solution, transition to next stage, re-estimate, and re-solve.
- SDDP versus MPC
  - Guan, Z., SDDP Production planning model, Fonterra report, 2019.
  - Martin, T., Stochastic optimization for the procurement of crude oil in refineries, PhD thesis, CERMICS, 2021.

Experiments on Fonterra production-inventory problem

[Z. Guan, 2019]

	Clairvoyant	MPC	SDDP	SDDP	SDDP
			6000 cuts	9000 cuts	12000 cuts
Annual revenue	100%	92%	79%	91%	90%

Revenues from out-of-sample simulations of inventory policies at Fonterra dairy cooperative.

Experiments on crude oil production-inventory problem

policy	Expert	Triplet	MPC	SDP_{esp}	SDP_{CVaR}	Suc-SDP	optimum
1^{st} crude	H2	B3	H4	L2	L2	H5	H5
2^{nd} crude	L2	H4	L2	H1	H1	L2	L2
3^{rd} crude	B5	L4	B1	B1	B1	B1	B1
margin ($.10^7$\$)	5.13	5.58	7.490	6.39	6.39	7.491	7.491
gap (to Expert)	0	8.9%	46.0%	24.6%	24.6%	46.0%	46.0%

Table 7.4: Operational margin and combination of crude oils generated by each policy for the historical scenario of December 2020

Reproduced from T. Martin, https://pastel.archives-ouvertes.fr/tel-03700627/document

Summary

1. Background
2. Certainty equivalence and myopic solutions
3. DRO and out-of-sample improvement
4. MPC and out-of-sample improvement

Summary

When does MPC = SDP?

- [Simon, 1956] Univariate control problem

$$x_{t+1} = x_t + u_t + \varepsilon_t$$

  quadratic costs on u and x, random ε_t.

- [Theil, 1957] Extensions to multivariate control problem

$$x_{t+1} = x_t + Bu_t + \varepsilon_t$$

- [Ziemba,1971] Extensions to constraints on controls.
- [See e.g. Bertsekas, 1976] Linear Quadratic Regulator

$$x_{t+1} = Ax_t + Bu_t + \varepsilon_t$$

- [Arrow et al, 1951, Bellman et al, 1955, Scarf, 1960] (S, s) inventory policies.
- [Mossin,1968] Multiperiod portfolio analysis.

Inventory example 1

Consider a company trading in a single product that incurs an inventory cost $R(x)$ for storage quantity x, where $R(x)$ is a differentiable, strictly convex increasing function with $R(0) = R'(0) = 0$. In each time period t the company observes the realization p of a random price P (i.i.d.) at which it can sell $u_t \leq x_t$.

Assumption (1)

u_t can be negative, allowing purchases at p.

Solution (c.f. [Bellman, 1955])

Optimal policy is myopic, i.e. sell/buy to achieve inventory target z where $R'(z) = (\mathbb{E}[P] - p)_+$ This is the same as optimal MPC policy (which optimizes using expected future prices).

Inventory example 2

Consider a company trading in a single product that incurs an inventory cost $R(x)$ for storage quantity x, where $R(x)$ is a differentiable, strictly convex increasing function with $R(0) = R'(0) = 0$. In each time period t the company observes the realization p of a random price P (i.i.d.) at which it can sell $u_t \leq x_t$.

Assumption (2)

$u_t \geq 0$, so purchases are not allowed.

Solution (DP)

Optimal DP policy stores to exploit option value in future prices.

Solution (MPC)

Optimal MPC policy (optimizing using expected future prices) stores less than DP.

Summary

Sample average approximation

Consider a stochastic optimization problem

$$\text{SP:} \quad \min_{x\in X} \mathbb{E}_{\mathbb{P}}\left[c(x,\xi)\right],$$

where ξ has probability measure $\mathbb{P}$. Given sample $S = \{\xi_1, \xi_2, \ldots, \xi_N\}$ the sample average approximation is

$$\text{SAA:} \quad \min_{x\in X} \mathbb{E}_{\mathbb{P}_0}\left[c(x,\xi)\right],$$

where $\mathbb{P}_0$ is the probability measure that assigns mass $\frac{1}{N}$ to each $\xi_i \in S$. The distributionally robust version of SAA sets $\mathcal{P}_\delta = \{\mathbb{Q} : d(\mathbb{Q}, \mathbb{P}_0) \leq \delta\}$ and solves

$$\text{DRO:} \quad \min_{x\in X} \sup_{\mathbb{Q}\in\mathcal{P}_\delta} \mathbb{E}_{\mathbb{Q}}\left[c(x,\xi)\right].$$

Sample average approximation

Assume that SP, SAA and DRO all have unique optimal solutions denoted x^*, $x_0(S)$, and $x_\delta(S)$ respectively. Observe that $x_0(S)$ and $x_\delta(S)$ depend on the sample S. For any $\delta \geq 0$, we let $\bar{x}_\delta = \mathbb{E}_S[x_\delta(S)]$. If $\bar{x}_\delta = x^*$ then the solution $x_\delta(S)$ is unbiased.

Definition

The out-of-sample improvement of $x_\delta(S)$ is

$$\mathbb{E}_S[\mathbb{E}_\mathbb{P}[c(x_0(S), \xi)]] - \mathbb{E}_S[\mathbb{E}_\mathbb{P}[c(x_\delta(S), \xi)]]$$

Bias and variance

Definition

The cost bias of $x_\delta(S)$ is

$$\beta_\delta = \mathbb{E}_{\mathbb{P}}[c(\bar{x}_\delta, \xi)] - \mathbb{E}_{\mathbb{P}}[c(x^*, \xi)].$$

Definition

The variance of $x_\delta(S)$ is

$$V_\delta = \mathbb{E}_S[(x_\delta(S) - \bar{x}_\delta)^2].$$

Example

Let $c(x, \xi(\omega)) = \frac{1}{2}x^2 - Z(\omega)x$ and $\delta = 0$. Then $x^* = \mathbb{E}_{\mathbb{P}}[Z(\omega)]$ and $x_0(S) = \mathbb{E}_{\mathbb{P}_0}[Z(\omega)]$. Since $\bar{x}_0 = \mathbb{E}_S[x_0(S)] = \mathbb{E}_{\mathbb{P}}[Z(\omega)]$ we have $\beta_0 = 0$.

Quadratic costs

Proposition

Suppose $c(x, \xi(\omega)) = \frac{1}{2}x^\top H(\omega)x^2 - Z(\omega)^\top x$. *A solution* $x_\delta(S)$ *to DRO yields a lower expected out-of-sample cost than a solution* $x_0(S)$ *to SAA if and only if*

$$\begin{aligned}\beta_\delta - \beta_0 \quad < \quad & \frac{1}{2}\mathbb{E}_S[(x_0(S) - \bar{x}_0)^\top \mathbb{E}[H](x_0(S) - \bar{x}_0)] \\ & - \frac{1}{2}\mathbb{E}_S[(x_\delta(S) - \bar{x}_\delta)^\top \mathbb{E}[H](x_\delta(S) - \bar{x}_\delta)].\end{aligned}$$

Bias-variance: LHS is the increase in cost bias and RHS measures a (scaled) reduction in variance of $x_\delta(S)$.

Corollary

If DRO reduces cost bias and variance then it gives positive out-of-sample improvement.

Quadratic example

Let $c(x, \xi(\omega)) = \frac{1}{2}x^2 - Z(\omega)x$, so $\beta_0 = 0$. Positive out-of-sample improvement $\iff$ $\beta_\delta < V_0 - V_\delta$.

Example (Anderson and P., 2021)

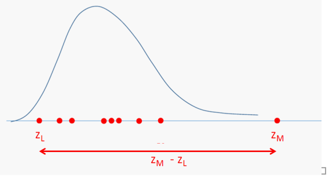
Sample $S = \{z_1, z_2, \ldots, z_N\}$ from $F(z)$, $\mathcal{P}_\delta = \{p : \sum_{i=1}^{N} \left|p_i - \frac{1}{N}\right| \leq \delta$.

Proposition

If F is symmetric then $\beta_\delta > V_0 - V_\delta$.

Proposition

If F is "right skewed" then $\beta_\delta < V_0 - V_\delta$ *for small enough* δ.



Summary

SAA for Inventory problem 2 (no purchases)

In Inventory problem 2, the sample average approximation yields a dynamic program DP using sample $S = \{p_1, p_2, \ldots, p_N\}$ where $\mathbb{P}_0$ is the probability measure that assigns mass $\frac{1}{N}$ to each $p_i \in S$ replacing P.
If purchasing is not allowed $(u_t \in \mathbb{R}_+)$ then:

Solution (SAA of DP)

For each price p_i there is an SAA inventory target $z_{SAA}(p_i)$. Sell $u^*(x, p_i) = (x - z_{SAA}(p_i))_+$.

Solution (MPC)

MPC inventory target $z_{MPC}(p_i)$ solves $R'(z) = (\mathbb{E}_{\mathbb{P}_0}[P] - p_i)_+$. Sell $u^*(x, p_i) = (x - z_{MPC}(p_i))_+$.

Proposition

$z_{MPC}(p_i) < z_{SAA}(p_i)$. Given S, the MPC solution performs worse (in-sample) than the DP solution.

Out-of-sample improvement

- If purchasing is not allowed ($u_t \in \mathbb{R}_+$) then MPC solution stores less than DP solution (it ignores optionality).
- MPC solution solves DRO for $\mathcal{P}_\delta = \{\mathbb{P}_0, \mathbb{D}_0\}$ where $\mathbb{D}_0$ assigns unit mass to $\mathbb{E}_{\mathbb{P}_0}[P]$.
- Interpret as total variation uncertainty set on supp$(\mathbb{P}_0)\cup$supp$(\mathbb{D}_0)$.
- Predict negative out-of-sample improvement from MPC when P has a symmetric distribution?
- Predict positive out-of-sample improvement from MPC when P has a "right-skew" distribution?

Simulations assuming no purchasing

Assumption

$R(x) = \frac{1}{2}x^2$, $x_0 = 50$, 10^5 *simulations with* $N = 2$.

```
julia> include("2price-sims.jl")

 Expected SDP objective: -2726.2414
 Expected MPC objective: -2725.1122
 Improvement: -1.1292 std error: 0.6187
```

P has uniform density on $[0, 100]$

```
julia> include("2price-sims.jl")

 Expected SDP objective: -148.8841
 Expected MPC objective: -175.0027
 Improvement: 26.1186 std error: 0.7106
```

P has lognormal density $(1, 1)$

Simulations assuming no purchasing

Assumption

$R(x) = \frac{1}{2}x^2$, $x_0 = 50$, 10^5 *simulations with* $N = 3$.

```
julia> include("3price-simsV2.jl")

 Expected SDP objective: -2780.3698
 Expected MPC objective: -2767.7264
 Improvement: -12.6434 std error: 0.6542
```

P has uniform density on $[0, 100]$

```
julia> include("3price-simsV2.jl")

 Expected SDP objective: -172.8223
 Expected MPC objective: -199.7452
 Improvement: 26.9228 std error: 0.6593
```

P has lognormal density $(1, 1)$

Conclusions

- Why is model predictive control very popular in practice?
  - DP is hard in high dimensions;
  - DP is hard if there are complicated constraints.
- SDDP can overcome these impediments, but requires finite distributions.
- SDDP applies SAA with modest sample size to stage problem.
- Using SAA the optimal solution to DP can give a worse solution than MPC when evaluated out of sample.
- Interpretation of MPC as distributionally robust optimization is fruitful.

References

- Arrow, K., Harris, T. and Marschak, J., Optimal inventory policy, *Econometrica*, 19, 250-272 (1951).
- Anderson, E.J. and Philpott, A.B., Improving sample average approximation using distributional robustness, *INFORMS Journal on Optimization*, 2021.
- Bellman, R, Glicksberg, I. and Gross, G. On the optimal inventory equation, *Management Science*, 1955.
- Bertsekas, D.P., *Dynamic programming and stochastic control*, Athena Scientific, 1976.
- Guan, Z., SDDP production planning model and SDDP hedge contract model, Report for Fonterra, 2019.
- Martin, T., Stochastic optimisation for the procurement of crude oil in refineries. General Mathematics [math.GM]. École des Ponts ParisTech, English. NNT : 2021ENPC0041. tel-03700627, 2021.

References

- Mossin, J., Optimal mutiperiod portfolio policies. *The Journal of Business*, 41(2), 1968.Scarf, H., The Optimality of (S, s) Policies in the Dynamic Inventory Problem, in Mathematical Methods in the Social Sciences, K. J. Arrow, S. Karlin, and P. Suppes (eds.), Stanford University Press, Stanford, 1960.
- Simon, H.A., Dynamic programming under uncertainty with a quadratic criterion function. *Econometrica,* 1956.
- Theil, H., A note on certainty equivalence in dynamic planning. *Econometrica*, 1957.