Multilevel Monte Carlo for VaR

Mike Giles, Wenhui Gou, Abdul-Lateef Haji-Ali

Mathematical Institute, University of Oxford
(BNP Paribas, Hong Kong)

(also discussions with Ralf Korn, Klaus Ritter)

Advances in Financial Mathematics

Paris, January 12, 2017

Outline

- MLMC and randomised MLMC
- Value-at-Risk and other risk measures
- prior research on VaR
  - Gordy & Juneja (2010)
  - Broadie, Du & Moallemi (2011)
- portfolio sub-sampling
- estimating inner conditional expectation
- adding in Euler-Maruyama or Milstein timestepping

Multilevel Monte Carlo

MLMC is based on the telescoping sum

$$\mathbb{E}[P_L] \;=\; \mathbb{E}[P_0] + \sum_{\ell=1}^{L} \mathbb{E}[P_\ell - P_{\ell-1}] \;\equiv\; \sum_{\ell=0}^{L} \mathbb{E}[\Delta P_\ell]$$

where P_ℓ represents an approximation of some output P on level ℓ, and $\Delta P_\ell \equiv P_\ell - P_{\ell-1}$ with $P_{-1} \equiv 0$.

If the weak convergence is

$$\mathbb{E}[P_\ell - P] = O(2^{-\alpha\,\ell}),$$

and Y_ℓ is an unbiased estimator for $\mathbb{E}[P_\ell - P_{\ell-1}]$, with variance

$$\mathbb{V}[Y_\ell] = O(2^{-\beta\,\ell}),$$

and expected cost

$$\mathbb{E}[C_\ell] = O(2^{\gamma\,\ell}), \quad \ldots$$

Multilevel Monte Carlo

... then the finest level L and the number of samples N_ℓ on each level can be chosen to achieve an RMS error of ε at an expected cost

$$C = \begin{cases} O\left(\varepsilon^{-2}\right), & \beta > \gamma, \\ O\left(\varepsilon^{-2}(\log \varepsilon)^2\right), & \beta = \gamma, \\ O\left(\varepsilon^{-2-(\gamma-\beta)/\alpha}\right), & 0 < \beta < \gamma. \end{cases}$$

I always try to get $\beta > \gamma$, so the main cost comes from the coarsest levels – use of QMC can then give substantial additional benefits.

With $\beta > \gamma$, can also randomise levels to eliminate bias (Rhee & Glynn, 2015).

Randomised Multilevel Monte Carlo

Starting from

$$\mathbb{E}[P] \;=\; \sum_{\ell=0}^{\infty} \mathbb{E}[\Delta P_\ell] \;=\; \sum_{\ell=0}^{\infty} p_\ell \; \mathbb{E}[\Delta P_\ell / p_\ell],$$

Rhee & Glynn's unbiased single-term estimator is

$$Y = \Delta P_{\ell'} \,/\, p_{\ell'}\,,$$

where ℓ' is a random integer which takes value ℓ with probability p_ℓ.

$\beta > \gamma$ is required to simultaneously obtain finite variance and finite expected cost using

$$p_\ell \propto 2^{-(\beta+\gamma)\ell/2}.$$

The complexity is then $O(\varepsilon^{-2})$.

Value-at-Risk

Financial institutions (banks, pension companies, insurance companies) hold portfolios with a variety of financial assets:

- cash
- bonds
- stocks
- options

and also debts / obligations:

- pension payments
- insurance payments

Value-at-Risk

Collectively, the portfolio value can be expressed as a sum of risk-neutral expectations of discounted payoffs/cash-flows f_p:

$$V = \sum_{p=1}^{P} \mathbb{E}[f_p]$$

in which the individual expectations are obtained in a variety of ways:

- actual value (e.g. cash and stocks)
- analytically (e.g. Black-Scholes option prices)
- quasi-analytically (highly efficient FFT methods)
- simple Monte Carlo
- complex Monte Carlo with time-stepping approximation of SDEs
- finite difference approximation of PDE

Value-at-Risk

The institutions, and the regulators, are concerned about the risk of a very large loss in a short time.

Given a risk horizon τ (1 week for banks, 1 year for pension / insurance companies?) with a given distribution for risk factors R_τ over that interval, the simplest question is

What is the probability of the portfolio loss L exceeding L_{max}?

This means estimating $\mathbb{P}[L > L_{max}] \equiv \mathbb{E}\left[\mathbf{1}(L > L_{max})\right]$ where

$$L(R_\tau) \;=\; \sum_{p=1}^{P} L_p(R_\tau) \;=\; \sum_{p=1}^{P} \mathbb{E}[f_p] - \mathbb{E}[f_p | R_\tau]$$

This is therefore a nested simulation problem, and the indicator function makes it even harder.

Value-at-Risk

The true VaR L_α is defined implicitly by

$$\mathbb{P}[L > L_\alpha] = \alpha$$

for some specified small α.

This involves either a root-finding process to determine L_α, or ordering multiple samples of L to find the appropriate quantile.

Another important risk measure is Conditional Value-at-Risk (CVaR), also known as Expected Shortfall,

$$\mathbb{E}\left[L \mid L > L_\alpha \right].$$

Value-at-Risk

What makes it expensive? Where is the potential for MLMC?

- large number of financial products in the portfolio (P)
- often needs lots of Monte Carlo samples for inner conditional expectation (M)
- sometimes needs lots of timesteps for SDE approximation (T)

P, M and T all offer possibilities for MLMC treatment

Prior research on VaR

Gordy & Juneja (2010) considered

$$\mathbb{P}\left[L > L_{max}\right] \;\equiv\; \mathbb{E}\left[\mathbf{1}\left(L > L_{max}\right)\right]$$

using N outer samples for R_τ, and M inner samples to estimate $L(R_\tau)$.

The variance for the estimator for $L(R_\tau)$ is $O(M^{-1})$, and Gordy & Juneja prove this produces a bias in the outer estimate of the same order.

Hence, for ε RMS accuracy require

- $M = O(\varepsilon^{-1})$
- $N = O(\varepsilon^{-2})$

and so the complexity is $O(M\,N\,P) = O(\varepsilon^{-3}P)$ since each inner sample has $O(P)$ cost.

Prior research on VaR

They also considered what happens as the number of products $P \to \infty$.

For this, they introduced a weighting $1/P$ for each product, so "total loss" is now "average loss".

In this case, the variance for the estimator for $L(R_\tau)$ is $O(M^{-1}P^{-1})$, if using independent sampling for each product.

Hence, for ε RMS accuracy require

- $M = \max(1, O(\varepsilon^{-1}P^{-1}))$
- $N = O(\varepsilon^{-2})$

and so the complexity is $O(M\,N\,P) = O(\max(\varepsilon^{-2}P, \varepsilon^{-3}))$.

Prior research on VaR

Their analysis can be generalised if we need to approximate an SDE: if the inner conditional expectation estimate has bias μ and variance σ^2, then overall the bias in the outer expectation is

$$O(\mu + \sigma^2).$$

Interesting – standard Mean Square Error analysis for SDE approximations without nested simulation gives

$$\text{MSE} = \mu^2 + \sigma^2$$

and we usually balance these two terms so that $\mu \sim \sigma \sim \varepsilon$.

However, this nested simulation application needs $\mu \sim \sigma^2 \sim \varepsilon$ so $\mu \ll \sigma$ – ideally we'd like it to be unbiased.

Prior research on VaR

Broadie, Du & Moallemi (2011) improved on Gordy & Juneja by noting that we don't need many samples to determine whether $L > L_{max}$ unless $L - L_{max}$ is small.

Heuristic analysis: when using M inner samples, if

$$\sigma^2(R_\tau) = \mathbb{V}[\Delta f \mid R_\tau], \qquad d(R_\tau) = |L - L_{max}|$$

where Δf is a single sample of the conditional loss, then usual confidence interval is $\pm 3\,\sigma/\sqrt{M}$ so need roughly

$$M = 9\,\sigma^2(R_\tau)/d^2(R_\tau)$$

inner samples to be sure whether or not $L > L_{max}$.

Prior research on VaR

Remembering $\sigma^2 \sim P^{-1}$ in the large P asymptotic analysis, if we use

$$M = \lceil \min\left(c\,\varepsilon^{-1} P^{-1}, 9\,\sigma^2(R_\tau)/d^2(R_\tau)\right) \rceil$$

then the cross-over point is at $d = O(\varepsilon^{1/2})$ and the average number of inner samples is

$$\overline{M} = \max(1, O(\varepsilon^{-1/2} P^{-1})),$$

reducing the overall complexity to $O(\overline{M}\,N\,P) = O(\max(\varepsilon^{-2} P, \varepsilon^{-5/2}))$.

This is better, but still not the $O(\varepsilon^{-2})$ that we aim for.

Also, the issue of timestepping approximation hasn't been addressed yet.

Wenhui Gou's MSc dissertation

- addressed large P issue
- considered simple application with Black-Scholes formula for inner conditional expectations
- approximated distribution of loss using Maximum Entropy reconstruction technique based on moments of loss $\phi(L)$
- developed control variate based on "delta-gamma" quadratic approximation

Wenhui Gou's MSc dissertation

Key idea: conditional on R_τ, the total loss is

$$\sum_{p=1}^{P} L_p = P \; \mathbb{E}[L_p]$$

where p is uniformly distributed in $\{1, 2, \ldots, P\}$ in the r.h.s. expectation

Hence, it can be approximated by

$$\sum_{p=1}^{P} L_p \approx \frac{P}{M} \sum_{m=1}^{M} L_{p_m}$$

with M i.i.d. indices p_m.

Wenhui Gou's MSc dissertation

Can then use $M_\ell = 2^\ell$ samples on level ℓ, with an antithetic estimator.

This means using an average over M_ℓ values p_m for "fine" level, and splitting these into two sets of $M_{\ell-1}$ values for two "coarse" estimates.

MLMC estimator for $\Delta\phi(L)$ on level ℓ is then

$$Y_\ell = \phi(L^{(f)}) - \frac{1}{2}\left(\phi(L^{(c,a)}) + \phi(L^{(c,b)})\right).$$

Analysis in G (2015) shows this results in

- bias $\sim 2^{-\ell}$
- variance $V_\ell \sim 4^{-\ell}$
- cost $C_\ell \sim 2^\ell$

so $\alpha \approx 1$, $\beta \approx 2$, $\gamma \approx 1 \implies$ complexity is $O(\varepsilon^{-2})$, independent of P.

Wenhui Gou's MSc dissertation

The variance of the estimator can be improved by noting that

$$L_p \equiv \mathbb{E}[f_p] - \mathbb{E}[f_p | R_\tau] \approx -\Delta S_\tau \frac{\partial \mathbb{E}[f_p]}{\partial S_0}$$

when τ is small, and the overall loss is approximately

$$-\Delta S_\tau \sum_{p=1}^{P} \frac{\partial \mathbb{E}[f_p]}{\partial S_0} \equiv -\Delta S_\tau \, \Delta$$

where Δ is the overall Delta for the portfolio, which is likely to be small. Hence,

$$L = -\Delta S_\tau \, \Delta + \sum_{p=1}^{P} \left(L_p + \Delta S_\tau \frac{\partial \mathbb{E}[f_p]}{\partial S_0} \right)$$

so $\Delta S_\tau \, \partial \mathbb{E}[f_p] / \partial S_0$ is used as the control variate.

(Full "delta-gamma" control variate add in next order terms.)

New ideas

1) extend Wenhui Gou's work to Monte Carlo estimation of conditional expectations, and probability of exceeding L_{max}:

$$\sum_{p=1}^{P} L_p \approx \frac{P}{M} \sum_{m=1}^{M} \Big(\, f_p(R_m, W_m) - f_p(R_\tau, W_m) \, \Big)$$

where W_m represents all of the random inputs needed for the conditional expectation, and R_m is the extra random inputs for the time interval $[0, \tau]$ needed for the time 0 valuation.

This essentially combines the P and M issues into one, controlled by M.

New ideas

2) If we use $M_\ell = 4^\ell$ then error in inner estimate is $O(M_\ell^{-1/2}) = O(2^{-\ell})$.

There is $O(2^{-\ell})$ probability of being within $O(2^{-\ell})$ of indicator step, producing an $O(1)$ value for MLMC estimator.

Hence, the MLMC variance is $V_\ell \sim 2^{-\ell}$.

Also,

$$\text{bias} \sim 4^{-\ell}, \quad C_\ell \sim 4^\ell,$$

so $\alpha \approx 2$, $\beta \approx 1$, $\gamma \approx 2$ and hence the complexity is $O(\varepsilon^{-5/2})$, independent of P.

New ideas

3) better to add in Broadie's adaptive ideas, and use something like

$$M_\ell(R_\tau) = \max\left(c_1\, 2^\ell, \min\left(c_2\, 4^\ell, 9\, \sigma^2(R_\tau)/d^2(R_\tau)\right)\right)$$

in which case we get

$$\text{bias} \sim 4^{-\ell}, \quad V_\ell \sim 2^{-\ell}, \quad C_\ell \sim 2^\ell,$$

so $\alpha \approx 2$, $\beta \approx 1$, $\gamma \approx 1$ and hence the complexity is roughly $O(\varepsilon^{-2})$.

4) again it is really important to use a control variate to reduce the variance of the MLMC estimator

New ideas

5) what about adding in time-stepping?

Originally, I thought this would be challenging, and may require Multi-Index Monte Carlo, but now I think it may not be too tough.

For the inner conditional expectation what we want is an unbiased unit-cost estimator.

In many cases, can use Rhee & Glynn's unbiased single-term estimator based on randomised MLMC – then analysis in 3) remains valid, since each single-term sample has $O(1)$ expected cost.

In other cases, can maybe use inner timestepping-MLMC to estimate conditional expectation, but we need to make the bias very small so that bias $= O(\text{variance})$.

Conclusions

- I think VaR may be a great new application area for MLMC
- so far, banks haven't been very interested in MLMC, perhaps because the savings have been modest – with VaR, I think the savings may be quite large
- I think nested MLMC may be the way to handle time-stepping
- there are other things I haven't discussed:
  - optimising for varying cost of different portfolio components
  - VaR, CVaR and other risk measures
- we should have numerical results for talk at Global Derivatives

Webpages:

http://people.maths.ox.ac.uk/gilesm/mlmc.html

http://people.maths.ox.ac.uk/gilesm/mlmc_community.html