Multilevel Monte Carlo methods: extra bits

Mike Giles

Mathematical Institute, University of Oxford

UNSW, April 18, 2023

Outline

The first hour covered the basics of the Multilevel Monte Carlo method.

These slides cover several important extensions:

- MLQMC (G, Waterhouse, 2009)
- Multi-Index Monte Carlo (Haji-Ali, Nobile & Tempone, 2016)
- randomised MLMC (Rhee & Glynn, 2015)
- Richardson-Romberg extrapolation (Lemaire & Pagès, 2017)

Relevant papers are on my MLMC Community webpage

MLQMC

One important way to improve the computational efficiency is to switch from Monte Carlo samples (based on iid random numbers) to Quasi-Monte Carlo samples (based on quasi-random numbers).

First paper on this was in 2009 with Ben Waterhouse, one of Ian Sloan's postdocs, as a result of a visit to UNSW in 2007 – complexity reduced from $O(\varepsilon^{-2})$ to roughly $O(\varepsilon^{-1})$ in the best cases.

Ian and Frances Kuo have contributed hugely to MLQMC for SPDE applications, including comprehensive numerical analysis.

MLQMC

Numerical algorithm (G, Waterhouse):

1. start with $L=0$
2. get an initial estimate for V_L using 32 random offsets and $N_L = 1$
3. while $\displaystyle\sum_{\ell=0}^{L} V_\ell > \varepsilon^2/2$, try to maximise variance reduction per unit cost by doubling N_ℓ on the level with largest $V_\ell \,/\, (C_\ell \; N_\ell)$
4. if $L<2$ or the bias estimate is greater than $\varepsilon/\sqrt{2}$, set $L := L+1$ and go back to step 2

Three other MLMC extensions

- Multi-Index Monte Carlo – Haji-Ali, Nobile, Tempone (2016)
  - important extension to MLMC approach, combining MLMC with sparse grid methods
- unbiased estimation – Rhee & Glynn (2015)
  - randomly selects the level for each sample
  - no bias, and finite expected cost and variance if $\beta > \gamma$
- Richardson-Romberg extrapolation – Lemaire & Pagès (2017)
  - reduces the weak error, and hence the number of levels required
  - particularly helpful when $\beta < \gamma$

MIMC – Multi-Index Monte Carlo

Standard "1D" MLMC truncates the telescoping sum

$$\mathbb{E}[P] \;=\sum_{\ell=0}^{\infty}\; \mathbb{E}[\Delta \widehat{P}_\ell]$$

where $\Delta \widehat{P}_\ell \equiv \widehat{P}_\ell - \widehat{P}_{\ell-1}$, with $\widehat{P}_{-1} \equiv 0$.

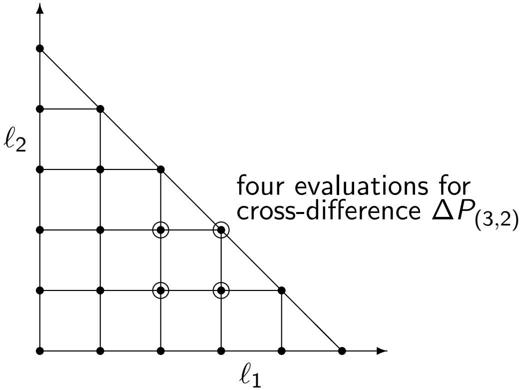
In "2D", Haji-Ali, Nobile & Tempone (2016) truncate the telescoping sum

$$\mathbb{E}[P] \;=\sum_{\ell_1=0}^{\infty}\;\sum_{\ell_2=0}^{\infty}\; \mathbb{E}[\Delta \widehat{P}_{\ell_1,\ell_2}]$$

where $\Delta \widehat{P}_{\ell_1,\ell_2} \equiv (\widehat{P}_{\ell_1,\ell_2} - \widehat{P}_{\ell_1-1,\ell_2}) - (\widehat{P}_{\ell_1,\ell_2-1} - \widehat{P}_{\ell_1-1,\ell_2-1})$

Different aspects of the discretisation vary in each "dimension" – for a 2D PDE, could use grid spacing $2^{-\ell_1}$ in direction 1, $2^{-\ell_2}$ in direction 2

MIMC – Multi-Index Monte Carlo



MIMC truncates the summation in a way which minimises the cost to achieve a target MSE – quite similar to sparse grids.

Can achieve $O(\varepsilon^{-2})$ complexity for a wider range of SPDE and other applications than plain MLMC.

Randomised MLMC

Rhee & Glynn (2015) start from

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

to develop an unbiased single-term estimator

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

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

$\beta > \gamma$ in the usual MLMC sense is required to simultaneously get finite variance and finite expected cost using

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

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

ML2R – Multilevel Richardson-Romberg

Lemaire & Pagès (2017) observed that if the weak error has expansion

$$\mathbb{E}[Y_\ell - Y] = \sum_{j=1}^{\infty} c_j \, 2^{-\alpha j \ell}$$

then can choose weights w_ℓ so that

$$\sum_{\ell=0}^{L} w_\ell \, \mathbb{E}[Y_\ell] = \mathbb{E}[Y] + R_L,$$

where $|R_L| = O(2^{-\alpha L(L+1)/2})$, much smaller than the usual MLMC bias which is $O(2^{-\alpha L})$.

ML2R – Multilevel Richardson-Romberg

To complete the ML2R formulation we need to set

$$\sum_{\ell=0}^{L} w_\ell \, \mathbb{E}[Y_\ell] = W_0 \, \mathbb{E}[Y_0] + \sum_{\ell=1}^{L} W_\ell \, \mathbb{E}[\Delta Y_\ell],$$

where

$$W_\ell \;=\; \sum_{\ell'=\ell}^{L} w_{\ell'} \;=\; 1 - \sum_{\ell'=0}^{\ell-1} w_{\ell'}.$$

Overall, very similar to MLMC but with weights for each level.

Asymptotically,

$$L_{\rm ML2R} = O(\sqrt{|\log \varepsilon|}\,),$$

instead of the usual

$$L_{\rm MLMC} = O(\,|\log \varepsilon|\,),$$

which gives big savings when $\beta < \gamma$.

Webpages

- research papers and talks:
  `people.maths.ox.ac.uk/gilesm/mlmc.html`
  `people.maths.ox.ac.uk/gilesm/slides.html`
- 70-page 2015 *Acta Numerica* review and MATLAB test codes:
  `people.maths.ox.ac.uk/gilesm/acta/`
- MATLAB and C++ software for lots of applications:
  `people.maths.ox.ac.uk/gilesm/mlmc/`
- community webpage listing groups and research papers using MLMC:
  `people.maths.ox.ac.uk/gilesm/mlmc_community.html`