Stochastic Numerical Analysis

Prof. Mike Giles

mike.giles@maths.ox.ac.uk

Oxford University Mathematical Institute

Inequalities

Stochastic numerical analysis requires a lot of tools, especially a surprisingly large collection of key inequalities.

- will introduce them one by one, with simple examples to show their usefulness
- in major applications later, the challenge will be seeing how to combine them appropriately to accomplish some task
- for the simple inequalities, I'll give proofs; for the complex ones I'll give references
- Wikipedia is a great reference source for the inequalities, and also has a lot of the proofs

Markov inequality

Given a positive scalar random variable X and a positive constant x,

$$\mathbb{P}[X \geq x] \;\leq\; \frac{\mathbb{E}[X]}{x}$$

Proof: let

$$f(X) = \begin{cases} 0, & X < x \\ x, & X \geq x \end{cases}$$

then $\quad X \geq f(X) \quad \Longrightarrow \quad \mathbb{E}[X] \geq \mathbb{E}[f(X)] = x\, \mathbb{P}[X \geq x]$

Corollary: (Chebyshev's inequality when $\mathbb{E}[X]\!=\!0$ and $p\!=\!2$) for a scalar r.v. X, a positive constant x, and $p\!\geq\!1$,

$$\mathbb{P}[\,|X| \geq x] \;\leq\; \frac{\mathbb{E}[\,|X|^p]}{x^p}$$

Proof: immediate given that $\mathbb{P}[\,|X| \geq x] = \mathbb{P}[\,|X|^p \geq x^p]$

Markov inequality

This is useful for bounding extreme behaviour.

For example, suppose $Z \sim N(0,1)$. Then, for any even integer p and $0 < \varepsilon \ll 1$, we have

$$\mathbb{P}[\,|Z| \geq \varepsilon^{-1}] \;\leq\; \varepsilon^p \,\mathbb{E}[\,Z^p]$$

which bounds the proportion of large values for Z.

Corollary: suppose $\Delta W \sim N(0,h)$, then for any $0\!<\!\delta\!\ll\!1$ we have

$$\mathbb{P}[\,|\Delta W| \geq h^{1/2-\delta}] \;\leq\; h^{p\delta}\,\mathbb{E}[\,|h^{-1/2}\Delta W|^p]$$

Markov inequality

In addition, if there are N Brownian increments then

$$\begin{aligned}\mathbb{P}[\max_n |\Delta W_n| \geq h^{1/2-\delta}] &\leq \sum_n \mathbb{P}[\,|\Delta W_n| \geq h^{1/2-\delta}] \\ &\leq N\, h^{p\delta}\, \mathbb{E}[\,|h^{-1/2}\Delta W|^p] \\ &\leq T\, h^{p\delta-1}\, \mathbb{E}[\,|h^{-1/2}\Delta W|^p]\end{aligned}$$

so if p is chosen so that $p\delta > 2$ then

$$\mathbb{P}[\max_n |\Delta W_n| \geq h^{1/2-\delta}] = o(h)$$

Extreme behaviour

When dealing with expected values, we need to be concerned about extreme behaviour if there is a possibility of these having extreme values.

e.g. what is $\mathbb{E}\left[\,\log_{10}(\text{human population in 2020})\right]$?

Extreme behaviour

More relevant example: consider SDE

$$\mathrm{d}S_t = -\, S_t^3 \,\mathrm{d}t + \mathrm{d}W_t$$

starting from $S_0 = 0$, and approximated using Euler method as

$$\widehat{S}_{n+1} = \widehat{S}_n - \widehat{S}_n^3 \, h + \Delta W_n.$$

with $h = 1/N$.

What is $\lim_{N\to\infty} \mathbb{E}[\widehat{S}_N^2]$?

Extreme behaviour

Putting

$$v_n = -\,(-1)^n\,\tfrac{1}{2}\,h^{1/2}\,\widehat{S}_n$$

we get

$$v_{n+1} = 4v_n^3 - v_n + (-1)^n\,\tfrac{1}{2}\,h^{1/2}\,\Delta W_n$$

If $v_n \geq 2$ and $|\Delta W_n| < h^{-1/2}$ then

$$v_{n+1} \;\geq\; v_n^3 \;>\; 2$$

so, by induction, if $v_1 \geq 2$ and $\max_n |\Delta W_n| < h^{-1/2}$ then

$$v_N \geq 2^{3^{N-1}}$$

a super-exponential blow-up.

Extreme behaviour

What is $\mathbb{P}[v_1 > 2]$? This requires $\Delta W_0 > 2\, h^{-1/2}$, and the probability of this is approximately

$$\frac{1}{\sqrt{8\,\pi}\ h}\ \exp\left(-2\, h^{-2}\right)$$

The probability of then having $\max_{n>0} |\Delta W_n| < h^{-1/2}$ is almost equal to 1 (using results from slide 5).

Hence, for sufficiently large N

$$\mathbb{E}[\widehat{S}_N^2] \ \geq\ \exp\left(-2\, N^2\right) \times 2^{3^{N-1}}$$

Note: $\log(\text{r.h.s.}) = -2\, N^2 + 3^{N-1} \log 2 \longrightarrow \infty$ as $N \to \infty$,
so $\mathbb{E}[\widehat{S}_N^2] \to \infty$ as $N \to \infty$

Extreme behaviour

This is a warning – can get extremely bad behaviour from samples with an extremely small probability.

In practice, probability is so small (in fact zero?) that you would be unlikely to see it experimentally for small h.

Nevertheless, it's a real problem for theoretical analysis.

Various ad-hoc fixes:

- "clamp" behaviour to limits of $\pm h^{-1}$
- clamp drift to $\pm h^{-1}$
- use adaptive time-stepping
- use drift-implicit approximation

Jensen inequality

This concerns convex functions $\phi(x)$, which satisfy

$$\phi\left((1-\lambda)\, x_1 + \lambda\, x_2\right) \;\leq\; (1-\lambda)\; \phi(x_1) + \lambda\; \phi(x_2)$$

for any $0 < \lambda < 1$.

There are multiple versions of Jensen's inequality:

- If $\sum_{i=1}^N \lambda_i = 1$ with $0<\lambda_i<1$ then

$$\phi\left(\sum_{i=1}^N \lambda_i\, x_i \right) \;\leq\; \sum_{i=1}^N \lambda_i\, \phi(x_i)$$

  Proof: by induction using
  $\sum_{i=1}^N \lambda_i\, x_i = (1-\lambda_N) \left(\sum_{i=1}^{N-1} \frac{\lambda_i}{1-\lambda_N}\, x_i \right) \;+\; \lambda_N x_N$

Jensen inequality

- taking $\lambda_i = N^{-1}$ gives

$$\phi\left(N^{-1}\sum_{i=1}^{N} x_i\right) \ \leq \ N^{-1}\sum_{i=1}^{N}\phi(x_i)$$

- for Riemann integrals, this becomes

$$\phi\left(\frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x\right) \ \leq \ \frac{1}{b-a}\int_a^b \phi(f(x))\,\mathrm{d}x$$

- finally, for expectations we get

$$\phi\left(\mathbb{E}[X]\right) \ \leq \ \mathbb{E}[\phi(X)]$$

Jensen inequality

The standard use is for $\phi(x) \equiv x^p$, for $p \geq 1$:

- $$\begin{aligned} \left| N^{-1} \sum_{i=1}^{N} x_i \right|^p &\leq N^{-1} \sum_{i=1}^{N} |x_i|^p \\ \Longrightarrow \quad \left| \sum_{i=1}^{N} x_i \right|^p &\leq N^{p-1} \sum_{i=1}^{N} |x_i|^p \end{aligned}$$

  this is sometimes used for small values of N (often 2-4) and sometimes for $N =$ # timesteps

Jensen inequality

- $$\left|\frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x\right|^p \;\leq\; \frac{1}{b-a}\int_a^b |f(x)|^p\,\mathrm{d}x$$

  $$\Longrightarrow \quad \left|\int_a^b f(x)\,\mathrm{d}x\right|^p \;\leq\; (b-a)^{p-1}\int_a^b |f(x)|^p\,\mathrm{d}x$$

- $$\big|\,\mathbb{E}[X]\,\big|^p \;\leq\; \mathbb{E}\left[\,|X|^p\,\right]$$

Hölder inequality

For any scalar random variables X, Y, and any $p, q > 1$ such that $1/p + 1/q = 1$,

$$\mathbb{E}[\,|XY|\,] \;\leq\; \mathbb{E}[\,|X|^p\,]^{1/p}\;\;\mathbb{E}[\,|Y|^q\,]^{1/q}$$

Proof: see Wikipedia – relies on Young's inequality for $a, b \geq 0$:

$$a\,b \;\leq\; \frac{a^p}{p} + \frac{b^q}{q}$$

Hölder inequality can be compared to the case in which X and Y are independent, which gives

$$\mathbb{E}[\,|XY|\,] \;=\; \mathbb{E}[\,|X|\,]\,\mathbb{E}[\,|Y|\,]$$

Hölder inequality

Hölder inequality is typically used for $p=q=2$ in which case it reduces to simply

$$\mathbb{E}[\,|XY|\,] \;\leq\; \sqrt{\mathbb{E}[\,|X|^2\,]\,\mathbb{E}[\,|Y|^2\,]}$$

which we've already proved by other means.

Sometimes used when $1/p=\delta,\ 1/q=1-\delta,\ \delta\ll 1$, giving

$$\mathbb{E}[\,|XY|\,] \;\leq\; \mathbb{E}[\,|X|^{1/\delta}\,]^{\delta}\;\; \mathbb{E}[\,|Y|^{1/(1-\delta)}\,]^{(1-\delta)}$$

Helpful when, for all p, $\mathbb{E}[\,|X|^p\,]$ is uniformly bounded for all h, and for all $q>1$, $\mathbb{E}[\,|Y|^q\,]=O(h^{1/2})$. Then, for any $\varepsilon>0$, we get

$$\mathbb{E}[\,|XY|\,]=o(h^{1/2-\varepsilon})$$

Hölder inequality

Another use: suppose

$$r = (1-\lambda)\, r_1 + \lambda\, r_2, \quad 0<\lambda<1.$$

then let $1/p = 1-\lambda,\ 1/q = \lambda$, and set

$$X = |Z|^{(1-\lambda)\, r_1}, \quad Y = |Z|^{\lambda\, r_2},$$

Hölder inequality then gives

$$\mathbb{E}[\, |Z|^r \,] \ \leq\ \mathbb{E}[\, |Z|^{r_1}]^{(1-\lambda)}\ \mathbb{E}[\, |Z|^{r_2}]^{\lambda}$$

so if $\mathbb{E}[\ |Z|^{r_1}\ ] = O(h^{\alpha_1})$, $\mathbb{E}[\ |Z|^{r_2}\ ] = O(h^{\alpha_2})$, then $\mathbb{E}[\ |Z|^r\ ] = O(h^{\alpha})$, with $\alpha = (1-\lambda)\, \alpha_1 + \lambda\, \alpha_2$

Grönwall inequality

Continuous versions:

- If $\beta(s)$ is continuous and non-negative and

$$u(t) \leq \alpha(t) + \int_0^t \beta(s)\, u(s)\, \mathrm{d}s$$

then

$$u(t) \leq \alpha(t) + \int_0^t \alpha(s)\, \beta(s) \exp\left(\int_s^t \beta(r)\, \mathrm{d}r \right) \, \mathrm{d}s$$

- when α, β are constants, it simplifies to

$$u(t) \ \leq \ \alpha \ \exp(\beta t)$$

Grönwall inequality

Discrete version: if β is non-negative and

$$u_n \leq \alpha + \sum_{j=0}^{n-1} \beta\, u_j$$

then
$$u_n \leq \alpha\ \exp(\beta n)$$

Proof by induction: if true for $0 \leq j < n$ then

$$\begin{aligned} u_n &\leq \alpha \left(1 + \sum_{j=0}^{n-1} \beta \exp(\beta j) \right) \\ &= \alpha \left(1 + \frac{\beta\,(\exp(\beta n) - 1)}{\exp(\beta) - 1} \right) \\ &\leq \alpha\ \exp(\beta n) \end{aligned}$$

Grönwall inequality

The Grönwall inequality turns out to be essential in the analysis of SDE approximations.

We will prove that the errors satisfy continuous integral (or discrete summation) inequalities, and then use the Grönwall inequality to turn these into bounds on the error.