SOME TITLE

Marius Jonsson

http://github.com/kingoslo/postmann.pat

May 23, 2017

ABSTRACT

This is a report submission for the first project of «Computational physics 2» at the Institute of Physics, University of Oslo, autumn 2016.

We prove a stronger version of the blocking method than Flydbjerg and Pettersen claim. We prove that if $X_1, X_2, \cdots, X_m$ are 2^n identically distributed random variables with stationary non-negative covariances, $\overline{X}_m = (1/m)\sum_{j=1}^m X_j$ denote their average and $X_i^{(j)}$ be a random variable subject to j blocking transformations. Then there exist a $k \in \mathbb{N}$ such that $k < \log_2 m$ and

$$\mathrm{Var}\left(\overline{X}_m\right) - \mathrm{Var}\left(X_i^{(k)}\right) < \frac{8}{m}\mathrm{Var}\left(X_i^{(k)}\right) \qquad \text{for all} \qquad 1 \leq i \leq 2^{-k}m.$$

INTRODUCTION

$$\psi_T(\mathbf{r}_1,\mathbf{r}_2) = C(\psi_0\psi_C)(\mathbf{r}_1,\mathbf{r}_2), \quad \psi_0(\mathbf{r}_1,\mathbf{r}_2) = \exp\left(-\frac{1}{2}\alpha\omega\left({r_1}^2+{r_2}^2\right)\right), \quad \psi_C(\mathbf{r}_1,\mathbf{r}_2) = \exp\left(\frac{\gamma r_{12}}{\beta r_{12}+1}\right)$$

$$E_T = \sum_{i=1}^N\left[-\frac{1}{2}\frac{\nabla_i^2\psi_T}{\psi_T} + \frac{1}{2}\omega^2 r_i^2 + \sum_{j=i+1}^N \frac{1}{r_{ij}}\right], \qquad \frac{\nabla_i^2\psi_T}{\psi_T} = \frac{\nabla_i^2\psi_0}{\psi_0} + \frac{\nabla_i^2\psi_C}{\psi_C} + 2\frac{\nabla_i\psi_C}{\psi_C}\frac{\nabla_i\psi_0}{\psi_0}$$

$$\frac{\nabla_i^2\psi_0}{\psi_0} = \alpha^2\omega^2 r_i^2 - 2\alpha\omega, \qquad \frac{\nabla_i^2\psi_C}{\psi_C} = \frac{\gamma(1+r_{ij}\gamma-\beta^2 r_{12}^2)}{r_{12}(1+\beta r_{12})^4}, \qquad \sum_{i=1}^2\frac{\nabla_i\psi_C}{\psi_C}\cdot\frac{\nabla_i\psi_0}{\psi_0} = -\frac{\alpha\omega\gamma r_{ij}}{(1+\beta r_{12})^2}$$

$$\frac{\nabla_i\psi_T}{\psi_T} = \mathbf{e}_1\left(-\alpha\omega x_i + \frac{\gamma(x_i-x_j)}{(1+\beta r_{ij})^2 r_{ij}}\right) + \mathbf{e}_2\left(-\alpha\omega y_i + \frac{\gamma(y_i-y_j)}{(1+\beta r_{ij})^2 r_{ij}}\right)$$

$$\frac{\partial}{\partial\theta_j}\langle E\rangle(\theta_1,\theta_2,\cdots\theta_n) = 2\left\langle\frac{E}{\psi_T}\frac{\partial\psi_T}{\partial\theta_j}\right\rangle - 2\left\langle\frac{1}{\psi_T}\frac{\partial\psi_T}{\partial\theta_j}\right\rangle\langle E\rangle$$

$$\frac{1}{\psi_T}\frac{\partial\psi_T}{\partial\alpha} = -\frac{\omega}{2}\left(r_1^2+r_2^2\right) \qquad \frac{1}{\psi_T}\frac{\partial\psi_T}{\partial\beta} = -\gamma\left(\frac{r_{12}}{1+\beta r_{12}}\right)^2$$

For more than two particles we generalize

$$\nabla\phi_{nx,ny} = \mathbf{e}_1\left(2n_x(\alpha\omega)^{1/2}\phi_{nx-1,ny} - \alpha\omega x\phi_{nx,ny}\right) + \mathbf{e}_2\left(2n_y(\alpha\omega)^{1/2}\phi_{nx,ny-1} - \alpha\omega y\phi_{nx,ny}\right)$$

$$\nabla^2\phi_{nx,ny} = 4\alpha\omega\left(n_x(n_x-1)\phi_{nx-2,ny} + n_y(n_y-1)\phi_{nx,ny-2}\right) - 4(\alpha\omega)^{3/2}\left(xn_x\phi_{nx-1,ny} + yn_y\phi_{nx,ny-1}\right) + \alpha\omega\phi_{nx,ny}\left(\alpha\omega r^2 - 2\right)$$

$$\frac{\nabla_k\psi_C}{\psi_C} = \sum_{i=1,\, i\neq k}^{n} \frac{\gamma_{ik}}{(1+\beta r_{ik})^2 r_{ik}}\Big(\mathbf{e}_1(x_k - x_i) + \mathbf{e}_2(y_k - y_i)\Big)$$

$$\frac{\nabla_i^2\psi_T}{\psi_T} = \frac{\nabla_i^2|D|_\uparrow}{|D|_\uparrow} + \frac{\nabla_i^2|D|_\downarrow}{|D|_\downarrow} + \frac{\nabla_i^2\psi_C}{\psi_C} + 2\left[\frac{\nabla_i|D|_\uparrow}{|D|_\uparrow} + \frac{\nabla_i|D|_\downarrow}{|D|_\downarrow}\right]\cdot\frac{\nabla_i\psi_C}{\psi_C}$$

$$\frac{\nabla_k^2\psi_C}{\psi_C} = \left[\sum_{j=1,j\neq k}^{n} \frac{\gamma_{kj}(x_k - x_j)}{(1+\beta r_{kj})^2 r_{kj}}\right]^2 + \left[\sum_{j=1,j\neq k}^{n} \frac{\gamma_{kj}(y_k - y_j)}{(1+\beta r_{kj})^2 r_{kj}}\right]^2 + \sum_{j=1,j\neq k}^{n} \frac{\gamma_{kj}(1-\beta r_{kj})}{r_{kj}(1+\beta r_{kj})^3}$$

The report is structured by «introduction»-, «methods»-, «results and discussion»- and finally a «conclusion and perspectives»-sections.

METHODS

$\varrho\varsigma\vartheta\varpi$

RESULTS AND DISCUSSION

Next we want to prove the blocking method. The paper by Flyvbjerg and Petersen (ref) gives an excellent introduction for applications. Mathematically interested readers may be affronted by the paper because the paper does not contain proof of the method. The paper falsely claims that the blocking transformations has a fixed point and that conditions can be put on γ_t such that repeated blockings ensure the $gamma_t$s wind up at $C\delta_{t0}$. This is false because the blocking transformations are not a linear operator (ref), and therefore T cannot have a fixed point (ref) moreover. Since there is no fixed point, there is no basin of attraction (ref). We will however prove the blocking method, and study its properties, such as convergence rate and compute the error induced by the algorithm and present sufficient conditions such that the method is stable and converges. To my knowledge, this has never been done for the blocking method. As you will see, the method is suitable for estimation of the variance of random variables.

Suppose $\overline{X}_m$ is the mean of m random variables and you want to compute $\mathrm{Var}\left(\overline{X}_m\right)$. The blocking method is a way of quickly computing $\mathrm{Var}\left(\overline{X}_m\right)$, and is implemented as follows: Start by taking the average of every two sequential random variables from a time-series. That means that we obtain $m/2$ new random variables. If we keep on repeating this process we will obtain $m/2^k$ random variables after k repetitions. We will call this a blocking transformation. It turns out that under certain conditions, there is some k, such that after k blocking transformations, the variance of the blocked variables are essentially the same as $\mathrm{Var}\left(\overline{X}_m\right)$. Let's make this precise.

Suppose that $X_1, \cdots, X_m$ are $m = 2^k$ identically distributed random variables. If we let $X_i^{(0)} = X_i$ for all $i \in \mathbb{N} = \{1, 2, \cdots\}$ and

$$x_i^{(k+1)} = \frac{1}{2}\left(x_{2i-1}^{(k)} + x_{2i}^{(k)}\right), \qquad \text{for all } k \in \mathbb{N}.$$

If $k \in \mathbb{N}$ and $m^{(k)} = 2^{-k}m$, we will say that the set $\{X_i^{(k)} \mid i \in \{1, 2, \cdots, m^{(k)}\}\}$ are ***subject to k blocking transformations***. We will later show that $m^{(k)} = 2^{-k}m$ is the appropriate number of random variables subject to k blocking transformations. Also, we will define the autocovariance $\gamma_{i,j} = \mathrm{Cov}(X_i, X_j)$. Snakk om stasjonær tidsserie og introduser γ_t.

PROPOSITION 1. *Suppose $X_1, \cdots, X_m$ are $m = 2^l$ identically distributed random variables with stationary non-negative covariances. Let $\overline{X}_m = (1/m)\sum_{j=1}^{m} X_j$ denote their average and $X_i^{(j)}$ be a random variable subject to j blocking transformations. Then*

$$\varepsilon_k = \mathrm{Var}\left(\overline{X}_m\right) - \mathrm{Var}\left(X_i^{(k)}\right) = \frac{2}{n^{(k)}} \sum_{t=1}^{n^{(k)}-1} \left(1 - \frac{t}{n^{(k)}}\right) (\gamma_k)_t \tag{1}$$

and

$$\gamma_{k,j} = \begin{cases} \frac{1}{2}\gamma_{k-1,2j} + \frac{1}{2}\gamma_{k-1,2j+1} & \textit{if } j = 0 \\ \frac{1}{4}\gamma_{k-1,2j-1} + \frac{1}{2}\gamma_{k-1,2j} + \frac{1}{4}\gamma_{k-1,2j+1} & \textit{if } j > 0 \end{cases} \tag{2}$$

Proof. See Flydbjerg and Petersen. A proof is also containted in the appendix. ■

LEMMA 1. *Suppose $m = 2^l$ denotes the number of observations and k denote the number of blocking transformations. Then*

$$k < \log_2 m, \qquad n^{(k)} = 2^{-k} m.$$

Proof. $n^{(k)} = 2^{-k}m$ and $n^{(k+1)} = 2^{-(k+1)}m$, and so

$$\frac{n^{(k)}}{n^{(k+1)}} = \frac{2^{-k}m}{2^{-(k+1)}m} = \frac{1}{2}$$

Rest to come ■

LEMMA 2. *Suppose γ_t denote the correlation between two observations X_i and X_j such that $|i - j| = t$. Suppose also that $\gamma_t \geq 0$ for all t. Then*

$$\gamma_t \leq \gamma_0 \qquad \textit{for all} \qquad t \geq 0.$$

Proof. We will need the formula $|\operatorname{Cov}(X, Y)|^2 \leq \operatorname{Cov}(X, X)\operatorname{Cov}(Y, Y)$ $(*)$ which is proven in the appendix. Write

$$\gamma_t^2 = |\gamma_t|^2 = |\gamma_{ij}|^2 = |\operatorname{Cov}(X_i, X_j)|^2 \overset{(*)}{\leq} \operatorname{Cov}(X_i, X_i)\operatorname{Cov}(X_j, X_j) = \gamma_{ii}\gamma_{jj} = \gamma_0^2$$

Since the function $f(\gamma) = \gamma^2$ is strictly increasing on $[0, \infty)$, we know $\gamma_t \geq \gamma_0$ follows. ■

PROPOSITION 2. *Suppose ε_k denotes the error after k blocking iterations. Let $\{\varepsilon_k\}_{k=1}^{d}$ denote the sequence of errors. Then this sequence is monotonously decreasing and the rate at which it decreases is*

$$\varepsilon_k - \varepsilon_{k+1} = \frac{(\gamma_k)_1}{n^{(k)}} \tag{3}$$

Proof. It remains to prove formula (3). We will manipulate sums and will be interested in which function $f(t)$ appear in the terms $\gamma_{f(t)}^{(k)}$ inside the summation. Define a functional $S_{f(j)} : \{2x - 1, 2x, 2x + 1\} \to \mathbb{R}$ by:

$$S_{f(j)} \equiv \frac{1}{n^{(k)}} \sum_{t=1}^{n^{(k)}/2-1} \left(1 - \frac{2t}{n^{(k)}}\right) \gamma_{f(t)}^{(k)}.$$

That means, using $n^{(k+1)} = n^{(k)}/2$, we can rewrite ε_{k+1} as

$$\begin{aligned} \varepsilon_{k+1} &\overset{(1)}{=} \frac{2}{n^{(k+1)}} \sum_{t=1}^{n^{(k+1)}-1} \left(1 - \frac{t}{n^{(k+1)}}\right) (\gamma_{k+1})_t \overset{(2)}{=} \frac{1}{n^{(k)}} \sum_{t=1}^{n^{(k)}/2-1} \left(1 - \frac{2t}{n^{(k)}}\right) \left(\gamma_{2t-1}^{(k)} + 2\gamma_{2t}^{(k)} + \gamma_{2t+1}^{(k)}\right) \\ &= S_{2j-1} + 2S_{2j} + S_{2j+1} = 2(S_{2j-1} + S_{2j}) + S_{2j+1} - S_{2j-1} \end{aligned} \tag{4}$$

where we added and subtracted S_{2j-1} in the last step. Let's examine the terms in the sum individually, write:

$$S_{2j-1} = \frac{1}{n^{(k)}} \sum_{t=1}^{n^{(k)}/2-1} \left(1 - \frac{2t}{n^{(k)}}\right) \gamma_{2t-1}^{(k)} = \frac{1}{n^{(k)}} \sum_{t=1}^{n^{(k)}/2-1} \left(1 - \frac{2t-1}{n^{(k)}}\right) \gamma_{2t-1}^{(k)} - \frac{1}{n^{(k)}} \sum_{t=1}^{n^{(k)}/2-1} \frac{1}{n^{(k)}} \gamma_{2t-1}^{(k)} \tag{5}$$

$$\begin{aligned} S_{2j+1} &= \frac{1}{n^{(k)}} \sum_{t=1}^{n^{(k)}/2-1} \left(1 - \frac{2t}{n^{(k)}}\right) \gamma_{2t+1}^{(k)} = \frac{1}{n^{(k)}} \left[\frac{2}{n^{(k)}} \gamma_3^{(k)} + \frac{4}{n^{(k)}} \gamma_5^{(k)} + \cdots + \frac{n^{(k)}-2}{n^{(k)}} \gamma_{n^k-1}^{(k)} \right] \\ &= \frac{1}{n^{(k)}} \sum_{t=2}^{n^k/2} \left(1 - \frac{2(t-1)}{n^{(k)}}\right) \gamma_{2t-1}^{(k)} = \frac{1}{n^{(k)}} \sum_{t=2}^{n^{(k)}/2} \left(1 - \frac{2t-1}{n^{(k)}}\right) \gamma_{2t-1}^{(k)} + \frac{1}{n^{(k)}} \sum_{t=2}^{n^{(k)}/2} \frac{1}{n^{(k)}} \gamma_{2t-1}^{(k)} \end{aligned}$$

Notice that if we subtract these terms, then a almost all terms from one of the sums cancel and we obtain

$$S_{2j+1} - S_{2j-1} = \frac{2}{n^{(k)}} \sum_{t=2}^{n^{(k)}/2} \frac{1}{n^{(k)}} \gamma_{2t-1}^{(k)} - \frac{1}{n^{(k)}} \left(1 - \frac{1}{n^{(k)}}\right) \gamma_1^{(k)} \tag{6}$$

Lets also investigate

$$\begin{aligned} n^{(k)}(S_{2j-1} + S_{2j}) &\overset{(5)}{=} \sum_{t=1}^{n^{(k)}/2-1} \left(1 - \frac{2t-1}{n^{(k)}}\right) \gamma_{2t-1}^{(k)} - \sum_{t=1}^{n^{(k)}/2-1} \frac{1}{n^{(k)}} \gamma_{2t-1}^{(k)} + \sum_{t=1}^{n^{(k)}/2-1} \left(1 - \frac{2t}{n^{(k)}}\right) \gamma_{2t}^{(k)} \\ &= \left[\left(1 + \frac{1}{n^{(k)}}\right) \gamma_1^{(k)} + \left(1 + \frac{2}{n^{(k)}}\right) \gamma_2^{(k)} + \cdots + \left(1 + \frac{n^{(k)}-2}{n^{(k)}}\right) \gamma_{n^{(k)}-2}^{(k)} \right] - \sum_{t=1}^{n^{(k)}/2-1} \frac{1}{n^{(k)}} \gamma_{2t-1}^{(k)} \\ &= \sum_{t=1}^{n^{(k)}-2} \left(1 - \frac{t}{n^{(k)}}\right) \gamma_t^{(k)} - \sum_{t=1}^{n^{(k)}/2-1} \frac{1}{n^{(k)}} \gamma_{2t-1}^{(k)} = \sum_{t=1}^{n^{(k)}-1} \left(1 - \frac{t}{n^{(k)}}\right) \gamma_t^{(k)} - \sum_{t=1}^{n^{(k)}/2} \frac{1}{n^{(k)}} \gamma_{2t-1}^{(k)} \end{aligned} \tag{7}$$

In the last equality we added and subtracted $(1 - t/n^{(k)})\gamma_t^{(k)}$ and in addition used that for $t = n^{(k)} - 1$, $(1 - t/n^{(k)})\gamma_t^{(k)} = 1/n^{(k)} \gamma_{n^{(k)}-1}^{(k)}$. This means that if we consider

$$2(S_{2j-1} + S_{2j}) \overset{(7)}{=} \frac{2}{n^{(k)}} \left[\sum_{t=1}^{n^{(k)}-1} \left(1 - \frac{t}{n^{(k)}}\right) \gamma_t^{(k)} - \sum_{t=1}^{n^{(k)}/2} \frac{1}{n^{(k)}} \gamma_{2t-1}^{(k)} \right] = \varepsilon_k - \frac{2}{n^{(k)}} \sum_{t=1}^{n^{(k)}/2} \frac{1}{n^{(k)}} \gamma_{2t-1}^{(k)} \tag{8}$$

Now substitute (6) and (8) into (4)

$$\begin{aligned} \varepsilon_{k+1} &\overset{(4)}{=} 2(S_{2j-1} + S_{2j}) + S_{2j+1} - S_{2j-1} \\ &\overset{(8)}{=} \varepsilon_k - \frac{2}{n^{(k)}} \sum_{t=1}^{n^{(k)}/2} \frac{1}{n^{(k)}} \gamma_{2t-1}^{(k)} + \frac{2}{n^{(k)}} \sum_{t=2}^{n^{(k)}/2} \frac{1}{n^{(k)}} \gamma_{2t-1}^{(k)} - \frac{1}{n^{(k)}} \left(1 - \frac{1}{n^{(k)}}\right) \gamma_1^{(k)} \\ &= \varepsilon_k - \frac{1}{n^{(k)}} \frac{\gamma^{(k}}{n^{(k)}} - \frac{\gamma_1^{(k)}}{n^{(k)}} + \frac{1}{n^{(k)}} \frac{\gamma^{(k}}{n^{(k)}} = \varepsilon_k - \frac{\gamma_1^{(k)}}{n^{(k)}} \end{aligned}$$

Subtract $\varepsilon_{k+1} - \gamma_1^{(k)}/n^{(k)}$ from each side of the equation, and the proposition follows. ■

LEMMA 3. *Suppose j and k are positive natural number and the sample size $m \geq 2^k(j-1) + 2^{k+1} - 1$, then $k < \log_2 m$ and*

$$\gamma_{k,j} = 2^{-2k}\Big[\gamma_{0,2^k(j-1)+1} + 2\gamma_{0,2^k(j-1)+2} + 3\gamma_{0,2^k(j-1)+3} + \cdots + 2^k\gamma_{0,2^k(j-1)+2^k}$$
$$+ (2^k-1)\gamma_{0,2^k(j-1)+2^k+1} + (2^k-2)\gamma_{0,2^k(j-1)+2^k+2} + \cdots + \gamma_{0,2^k(j-1)+2^{k+1}-1}\Big]. \quad (9)$$

Proof. We first show that $k < \log_2 m$. Fix j and k such that $m \geq 2^k(j-1) + 2^{k+1} - 1$, then

$$m \geq \underbrace{2^k(j-1)}_{\geq 0} + \underbrace{2^{k+1}-1}_{\geq 2^k+1} \geq 2^k+1 \qquad \text{only if} \qquad \log_2 m \geq \log_2(2^k+1) > \log_2 2^k = k\log_2 2 = k$$

We prove the rest of the lemma by induction. Fix j such that $m \geq 2^1(j-1) + 2^{1+1} - 1$, in particular this ensures that if $k = 1$, then $m \geq 2j + 1$ and therefore $\gamma_{0,2j+1}$ exists. Define $M = \sup_{k \in \mathcal{N}}\{m \geq 2^k(j-1) + 2^{k+1} - 1\}$. Assume $k = 1$ and write

$$\gamma_{1,j} \overset{(2)}{=} 2^{-2}\left(\gamma_{0,2j-1} + \gamma_{0,2j} + \gamma_{0,2j+1}\right)$$

Assume now that there exist a positive natural number $k < M$ such that equation (9) is true. This implies $k + 1 \leq M$ and hence $\gamma_{0,2^k2j+2^k+2^k-1}$ exists, and we can write

$$\gamma_{k+1,j} \overset{(2)}{=} 2^{-2}(\gamma_{k,2j-1} + 2\gamma_{k,2j} + \gamma_{k,2j+1}) \overset{(9)}{=}$$
$$2^{-2}2^{-2k}\Big(\gamma_{0,2^k(2j-2)+1} + 2\gamma_{0,2^k(2j-2)+2} + \cdots + 2^k\gamma_{0,2^k(2j-2)+2^k}$$
$$+ (2^k-1)\gamma_{0,2^k(2j-2)+2^k+1} + \cdots + \gamma_{0,2^k(2j-2)+2^k+2^k-1}\Big) +$$
$$2^{-2}2^{-2k}\Big(2\gamma_{0,2^k(2j-1)+1} + 4\gamma_{0,2^k(2j-1)+2} + \cdots + 2\,2^k\gamma_{0,2^k(2j-1)+2^k}$$
$$+ 2(2^k-1)\gamma_{0,2^k(2j-1)+2^k+1} + \cdots + 2\gamma_{0,2^k(2j-1)+2^k+2^k-1}\Big) +$$
$$2^{-2}2^{-2k}\Big(\gamma_{0,2^k2j+1} + 2\gamma_{0,2^k2j+2} + \cdots + 2^k\gamma_{0,2^k2j+2^k}$$
$$+ (2^k-1)\gamma_{0,2^k2j+2^k+1} + \cdots + \gamma_{0,2^k2j+2^k+2^k-1}\Big).$$

By using that $2^{-2}2^{-2k} = 2^{-2(k+1)}$ and factoring $\gamma_{0,K}$ together for all $2^{k+1}(j-1) + 1 \leq K \leq 2^{k+1}(j-1) + 2^{k+2} - 1$, the lemma follows. ■

PROPOSITION 3. *Suppose $\boldsymbol{\gamma}_k = \left((\gamma_k)_0, (\gamma_k)_1, \cdots, (\gamma_k)_{n_k}\right)$ denote the vector of correlation structure at blocking iteration number k. Suppose $m \geq 2^{k+1} - 1$, then $k < \log_2 m$ and,*

$$2^{2k}(\gamma_k)_1 = (\gamma_0)_1 + 2(\gamma_0)_2 + 3(\gamma_0)_3 + \cdots + 2^k(\gamma_0)_{2^k} + (2^k-1)(\gamma_0)_{2^k+1} + \cdots + (\gamma_0)_{2^{k+1}-1}$$

Proof. Use the previous lemma with $j = 1$. ■

Add def av stationary non negative decreasing covariance.

THEOREM 1 (Blocking method)**.** *Suppose $X_1, \cdots, X_m$ are $m = 2^l$ identically distributed random variables with stationary non-negative covariances. Let $\overline{X}_m = (1/m)\sum_{j=1}^m X_j$ denotes their average and $X_i^{(j)}$ be a random variable subject to j blocking transformations. Then there exist a $k \in \mathbb{N}$ such that*

$$k < \log_2 m \qquad \text{and} \qquad \mathrm{Var}\left(\overline{X}_m\right) - \mathrm{Var}\left(X_i^{(k)}\right) < \frac{8}{m}\mathrm{Var}\left(X_i^{(k)}\right) \qquad \text{for all} \qquad 1 \leq i \leq 2^{-k}m.$$

Proof. To come. ■

COROLLARY 1. *Suppose $X_1, \cdots, X_m$ are $m = 2^l$ are observations from a stationary time series. Let $\overline{X}_m = (1/m)\sum_{j=1}^{m} X_j$ denote their average and $X_i^{(j)}$ be a random variable subject to j blocking transformations. Then there exist a $k \in \mathbb{N}$ such that*

$$k < \log_2 m \qquad \text{and} \qquad \mathrm{Var}\left(\overline{X}_m\right) - \mathrm{Var}\left(X_i^{(k)}\right) < \frac{8}{m}\mathrm{Var}\left(X_i^{(k)}\right) \qquad \text{for all} \qquad 1 \leq i \leq 2^{-k}m.$$

Proof. To do vis at positiviteten til gammaene er implisitt ■

COROLLARY 2. *Suppose $X_1, \cdots, X_m$ are $m = 2^l$ are observations from a stationary Markov chain. Let $\overline{X}_m = (1/m)\sum_{j=1}^{m} X_j$ denote their average and $X_i^{(j)}$ be a random variable subject to j blocking transformations. Then there exist a $k \in \mathbb{N}$ such that*

$$k < \log_2 m \qquad \text{and} \qquad \mathrm{Var}\left(\overline{X}_m\right) - \mathrm{Var}\left(X_i^{(k)}\right) < \frac{8}{m}\mathrm{Var}\left(X_i^{(k)}\right) \qquad \text{for all} \qquad 1 \leq i \leq 2^{-k}m.$$

Proof. To do ■

Moreover $S^2 = blabla$ is an unbiased estimator of $\mathrm{Var}\left(\overline{X}_m\right)$. Bla bla.
This theorem says that after less than $\log_2 m$ blocking iterations, the difference between the variance of the mean and variance of the blocked variables, is essentially zero. This is the case since m is typically many orders of magnitude larger than $\mathrm{Var}\left(X_i^{(k)}\right)$.

CONCLUSION AND PERSPECTIVES

APPENDIX

In this section we will prove the Hastings-Metropolis theorem. Often whilst proving results about Markov chains we are interested in whether the Markov chain can reach every state from any other state. In the following proof this is important because it ensures that all the divisions we will make are non-zero. Let's make this notion precise. Suppose q_{ij} is the transition probability matrix of a Markov chain X_n and S is the state space. We say that X_n is ***irreducible*** if for all $i, j \in S$ there exist an $n \in \mathbb{N}$ such that $(q^n)_{ij}, (q^n)_{ji} > 0$.

Suppose that X_n is irreducible and not deterministically periodic, then we say that the probabilities $\pi_i = P(X_n = i)$ is the ***stationary distribution*** of X_n.

We will say that X_n is ***time reversible*** if the conditional probability $P(X_n = j, \text{given that } X_{n+1} = i) = P_{ij}$ for all $i, j \in S$ and $n \in \mathbb{N}$. Interestingly, it is easily shown that a Markov chain is time reversible if and only if $\pi_i P_{ij} = \pi_j P_{ji}$ for all $i, j \in S$. If you wish to prove this yourself, the proposition follows from Markov chain property and Bayes theorem. With these definitions, we are ready for the main result.

THEOREM (Hastings-Metropolis theorem)**.** *Suppose that $C\pi_i$ is a discrete probability distribution. If q_{ij} is any irreducible transition probability matrix, and X_n is a Markov chain with transition probability matrix*

$$P_{ij} = \begin{cases} \alpha_{ij} q_{ij}, & j \neq i \\ q_{ii} + \sum_{k=0}^{\infty} q_{ik}(1 - \alpha_{ik}) & j = i \end{cases}, \qquad \text{where} \qquad \alpha_{ij} = \min\left(\frac{\pi_j q_{ji}}{\pi_i q_{ij}}, 1\right), \tag{10}$$

then X_n is time reversible with stationary distribution π_i.

Proof. Assume that the hypothesis is true. Then in particular $q_{ij} \neq 0$ for all i, j since q is irreducible. Notice that if

$$\frac{\pi_j q_{ji}}{\pi_i q_{ij}} = 1,$$

then there is nothing to prove since then $\alpha_{ij} = 1$ and $\alpha_{ji} = 1$, and therefore

$$\pi_i P_{ij} = \pi_j P_{ji} \tag{11}$$

is automatic. Hence it suffices to prove (11) for the two cases

$$\frac{\pi_j q_{ji}}{\pi_i q_{ij}} > 1 \qquad \text{and} \qquad \frac{\pi_j q_{ji}}{\pi_i q_{ij}} < 1,$$

separately. Suppose first that $\pi_j q_{ji} > \pi_i q_{ij}$ (†). Write:

$$\pi_i P_{ij} \overset{(10)}{=} \pi_i q_{ij} \alpha_{ij} \overset{(10)(\dagger)}{=} \pi_i q_{ij} \cdot 1 = \pi_i q_{ij} \frac{\alpha_{ji}}{\alpha_{ji}} = \alpha_{ji} \pi_i q_{ij} \frac{1}{\alpha_{ji}} \overset{(\dagger)}{=} \alpha_{ji} \pi_i q_{ij} \frac{\pi_j q_{ji}}{\pi_i q_{ij}} = \alpha_{ji} \pi_j q_{ji} \overset{(10)}{=} \pi_j P_{ji}.$$

In the case that $\pi_j q_{ji} < \pi_i q_{ij}$ (‡), write

$$\pi_i P_{ij} \overset{(10)}{=} \pi_i q_{ij} \alpha_{ij} \overset{(10)(\ddagger)}{=} \pi_i q_{ij} \frac{\pi_j q_{ji}}{\pi_i q_{ij}} = \pi_j q_{ji} = \pi_j q_{ji} \cdot 1 \overset{(10)(\ddagger)}{=} \pi_j q_{ji} \cdot \alpha_{ji} = \pi_j P_{ji},$$

which means X_n is time reversible with stationary probability π_i. ∎

PROPOSITION. *Suppose $X_1, \cdots, X_m$ are $m = 2^l$ identically distributed random variables with stationary non-negative covariances. Let $\overline{X}_m = (1/m) \sum_{j=1}^m X_j$ denote their average and $X_i^{(j)}$ be a random variable subject to j blocking transformations. Then*

$$\varepsilon_k = \operatorname{Var}\left(\overline{X}_m\right) - \operatorname{Var}\left(X_i^{(k)}\right) = 2^{k+1} \sum_{t=1}^{2^{-k}m-1} \left(1 - \frac{2^k t}{m}\right) \frac{(\gamma_k)_t}{m}$$

This is not a new result. This was shown by ffFlydbjerg og company. The proof is however containted in the appendix for completness.

Proof. Define the mean $\overline{X}_n = (1/n) \sum_{k=1}^n X_n$. Let's compute the variance of $\overline{X}_n$. Write

$$\operatorname{Var}\left(\overline{X}_n\right) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \left\langle X_i X_j \right\rangle - \left\langle X_i \right\rangle \left\langle X_j \right\rangle = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \operatorname{Cov}(X_i, X_j) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \Sigma_{ij}, \tag{12}$$

where $\Sigma_{ij} \equiv \operatorname{Cov}(X_i, X_j)$ are the elements of the covariance matrix Σ. In particular, that means that $\Sigma_{ii} = \operatorname{Cov}(X_i, X_i) = \operatorname{Var}(X_i) = \sigma^2$ (∗) and $\Sigma_{ij} = \operatorname{Cov}(X_i, X_j) = \operatorname{Cov}(X_j, X_i) = \Sigma_{ji}$ (†). Then hence by first splitting the summation at the right hand side of equation (12) over the diagonal of Σ and triangular parts of Σ separately, we obtain:

$$\begin{aligned}
\operatorname{Var}\left(\overline{X}_n\right) &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \Sigma_{ij} = \frac{1}{n^2} \Big(\sum_{i=1}^n \underbrace{\Sigma_{ii}}_{=\sigma^2\ (*)} + \sum_{i=1}^n \sum_{j=i+1}^n \Sigma_{ij} + \sum_{i=j+1}^n \sum_{j=1}^n \underbrace{\Sigma_{ij}}_{=\Sigma_{ji}\ (\dagger)} \Big) \\
&= \frac{1}{n^2} \Big(\underbrace{\sum_{i=1}^n \sigma}_{=n\sigma^2} + \sum_{i=1}^n \sum_{j=i+1}^n \Sigma_{ij} + \underbrace{\sum_{i=1}^n \sum_{j=i+1}^n \Sigma_{ji}}_{\sum_{j=i+1}^n \sum_{i=1}^n \Sigma_{ij}} \Big) = \frac{1}{n^2} \Big(n\sigma^2 + 2 \sum_{i=1}^n \sum_{j=i+1}^n \Sigma_{ij} \Big)
\end{aligned} \tag{13}$$

Since the observations are identically distributed the covariances satisfy $\Sigma_{ij} = \Sigma_{(i+k)(j+k)}$ for all $i, j, k \in \mathbb{N}$ (check this). That means that along diagonals in Σ, the components of Σ are constant.

$$\operatorname{Var}\left(\overline{X}_n\right) \overset{(13)}{=} \frac{1}{n} \Big(\sigma^2 + 2 \sum_{i=1}^n \sum_{j=i+1}^n \frac{1}{n} \Sigma_{ij} \Big) = \frac{1}{n} \Big(\sigma^2 + 2 \sum_{t=1}^{n-1} \sum_{j=i+1}^n \frac{1}{n} \Sigma_{ij} \Big)$$

Let's rewrite the last term. As you've seen, the sum is over the upper triangular (without the diagonal) elements of Σ. ∎

LEMMA. *Suppose X and Y are random variables with finite variance, then*

$$|\operatorname{Cov}(X,Y)|^2 \leq \operatorname{Cov}(X,X)\operatorname{Cov}(Y,Y)$$

Proof. Suppose Ω denotes the vector space of all real random variables with finite variance over the real numbers. Suppose $X, Y \in \Omega$ and define

$$(X,Y) = \langle XY \rangle\,. \tag{14}$$

To see that $\big(\Omega, (\cdot,\cdot)\big)$ is an inner product space, note that

$$\begin{aligned}
\text{Positivity:}\quad & (X,X) = \left\langle X^2 \right\rangle \geq 0 \qquad = 0 \quad \text{if and only if} \quad X = 0 \quad \text{almost surely.}\\
\text{Symmetry:}\quad & (X,Y) = \langle XY \rangle = \langle YX \rangle = (Y,X)\\
\text{Bilinearity:}\quad & (aX+bY,Z) = \langle (aX+bY)Z \rangle = a\,\langle XZ \rangle + b\,\langle YZ \rangle = a(X,Z) + b(YZ)
\end{aligned}$$

This proves that $\langle\cdot,\cdot\rangle$ is an inner product on Ω (ref mcdonald and weiss). In particular, that means we can use Cauchy-Schwarz (ref). Now use the definition of covariance $(*)$ and apply Cauchy-Schwarz inequality $(**)$. Write

$$\begin{aligned}
|\operatorname{Cov}(X,Y)|^2 &\overset{(*)}{=} |\,\langle (X-\langle X \rangle)(Y-\langle Y \rangle) \rangle\,|^2 \qquad 0 \overset{(14)}{=} |(X-\langle X \rangle\,, Y-\langle Y \rangle)|^2\\
&\overset{(**)}{\leq} (X-\langle X \rangle\,, X-\langle X \rangle)(X-\langle Y \rangle\,, Y-\langle Y \rangle) \overset{(*)(14)}{=} \operatorname{Cov}(X,X)\operatorname{Cov}(Y,Y)
\end{aligned}$$

Which is what we wanted. ■