Control Variates

Gersende FORT, Eric MOULINES

LTCI, CNRS / TELECOM ParisTech

Première partie I

Introduction

Principles

The principle of control variates consists in using the estimation error of known quantities, to improve the estimation of unknown quantities.
Assumptions

- X and Y are positively correlated.
- $\mathbb{E}[X]$ is known and $\{X_k, k \geq 0\}$ are i.i.d. with the same distribution than X.
- $\mathbb{E}[Y]$ is unknown and $\{Y_k, k \geq 0\}$ are i.i.d. with the same distribution than Y.

Principles

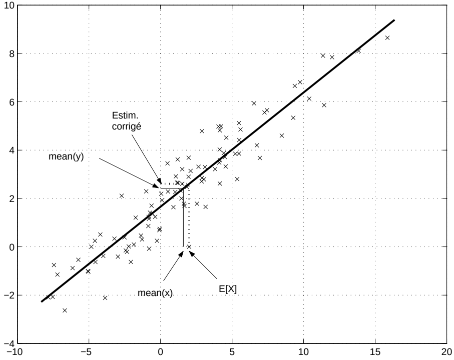
- The plain Monte Carlo estimators of $\mathbb{E}[Y]$ (resp. $\mathbb{E}[X]$) are given by

$$\bar{\mu}_n \stackrel{\text{def}}{=} \frac{1}{n}\sum_{k=1}^{n} Y_k \qquad\qquad \frac{1}{n}\sum_{k=1}^{n} X_k$$

- The estimation error for $\mathbb{E}[X]$ is

$$\mathcal{E}_n \stackrel{\text{def}}{=} \mathbb{E}[X] - \frac{1}{n}\sum_{k=1}^{n} X_k$$

- Idea : correct $\hat{\mu}_n$ knowing $\mathcal{E}_n$.


Estim.
corrigé
mean(y)
mean(x)
E[X]
10
8
6
4
2
0
−2
−4
−10
−5
0
5
10
15
20

If $\mathcal{E}_n > 0$,

- $n^{-1} \sum_{k=1}^{n} X_k$ underestimate $\mathbb{E}[X]$.
- Since X and Y are positively correlated, it is likely that $\bar{\mu}_n$ also underestimate $\mathbb{E}[Y]$
- idea Correct the estimator by adding a positive quantity

$$\bar{\mu}_n + b \left(\mathbb{E}[X] - \frac{1}{n} \sum_{k=1}^{n} X_k \right)$$

Control Variates estimators

- Estimator

$$\hat{\mu}_n(b) \stackrel{\text{def}}{=} \frac{1}{n}\sum_{k=1}^{n} Y_k - b\left(\frac{1}{n}\sum_{k=1}^{n} X_k - \mathbb{E}[X]\right)$$

- Questions :
  1. how to adjust b ?
  2. What are the properties of this estimator ? (how much do I gain on the variance)
- More general schemes (involving non linear corrections) may be considered as well.

Deuxième partie II

Control Variates : Methodology

Assumptions

1. $\{(X_k, Y_k), k \geq 0\}$ are i.i.d.
2. $\mathbb{E}[X]$ is known and Var$(X) < +\infty$.
3. Var$(Y) < +\infty$.

Beware

- $\{X_k, k \geq 0\}$ are i.i.d.
- $\{Y_k, k \geq 0\}$ are i.i.d.
- But X_k and Y_k are dependent !

Definition

- Plain MC estimator

$$\bar{\mu}_n \stackrel{\text{def}}{=} \frac{1}{n}\sum_{k=1}^{n} Y_k$$

- Control Variate estimator

$$\hat{\mu}_n(b) \stackrel{\text{def}}{=} \bar{\mu}_n - b \ \left(\frac{1}{n}\sum_{k=1}^{n} X_k - \mathbb{E}[X]\right) = \bar{\mu}_n - b \ \left(\bar{X}_n - \mathbb{E}[X]\right)$$

where

$$\bar{X}_n \stackrel{\text{def}}{=} \frac{1}{n}\sum_{k=1}^{n} X_k$$

Some easy properties

The estimator $\hat{\mu}_n(b)$ is

- unbiased (to show)

Some easy properties

The estimator $\hat{\mu}_n(b)$ is

- unbiased (to show)
- (strongly) consistent (to show)

Variance

- Variance (to show)

$$\mathrm{Var}\left(\hat{\mu}_n(b)\right) = \mathrm{Var}\left(\bar{\mu}_n\right) + \frac{1}{n}\Big(\mathrm{Var}(bX) - 2\mathrm{Cov}(bX, Y)\Big)$$

- An easy upper bound

$$\mathrm{Var}\left(\hat{\mu}_n(b)\right) \leq \mathrm{Var}\left(\bar{\mu}_n\right) \Longleftrightarrow \mathrm{Var}(bX) < 2\mathrm{Cov}(bX, Y).$$

- Idea : Choose b to minimize the variance $\mathrm{Var}\left(\hat{\mu}_n(b)\right)$.

Optimal b

- Optimal b (to show)

$$b_\star \stackrel{\text{def}}{=} \frac{\text{Cov}(X,Y)}{\text{Var}(X)}.$$

- Optimal Variance (to show)

$$\text{Var}\left(\hat{\mu}_n(b_\star)\right) = \text{Var}(\bar{\mu}_n)\left(1 - \rho_{X,Y}^2\right)$$

where

$$\rho_{X,Y} = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \text{Corr}(X,Y)$$

Limiting distribution

- Central Limit Theorem

$$\hat{\mu}_n(b) \stackrel{\text{def}}{=} \bar{\mu}_n - b \ \left(\frac{1}{n}\sum_{k=1}^{n} X_k - \mathbb{E}[X]\right)$$
$$= \mathbb{E}[Y] + \frac{1}{n}\sum_{k=1}^{n} \left(\{Y_k - \mathbb{E}[Y]\} - b\ \{X_k - \mathbb{E}[X]\}\right)$$

- Confidence bound

$$\frac{\sqrt{n}}{\sigma(b)}\left(\hat{\mu}_n(b) - \mathbb{E}[Y]\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

where

$$\sigma^2(b) \stackrel{\text{def}}{=} \mathsf{Var}(Y - bX) = \mathsf{Var}(Y) + b^2\mathsf{Var}(X) - 2b\mathsf{Cov}(X, Y).$$

Confidence Bounds

- The limiting distribution can be used to construct asymptotic confidence intervals with coverage $(1-\delta)\%$ of $\mathbb{E}[Y]$

$$\left[\hat{\mu}_n(b) - \frac{\sigma(b)}{\sqrt{n}}\ z_{1-\delta/2};\ \hat{\mu}_n(b) + \frac{\sigma(b)}{\sqrt{n}}\ z_{1-\delta/2}\right]$$

where z_α is the α-th quantile of the standard normal distribution.

Confidence Bounds

- The limiting distribution can be used to construct asymptotic confidence intervals with coverage $(1-\delta)\%$ of $\mathbb{E}[Y]$

$$\left[\hat{\mu}_n(b) - \frac{\hat{s}_n(b)}{\sqrt{n}}\ z_{1-\delta/2};\ \hat{\mu}_n(b) + \frac{\hat{s}_n(b)}{\sqrt{n}}\ z_{1-\delta/2}\right]$$

  where z_α is the α-th quantile of the standard normal distribution.

Wrap up

- The variance reduction increases with the correlation.
- The variance reduction vanishes whenever X and Y are independent.
- When $\rho^2_{X,Y} = 1$, the variance is zero ! !
- Problem make the dream comes true !

How to compute $b_\star$

$$b_\star \stackrel{\text{def}}{=} \frac{\mathsf{Cov}(X, Y)}{\mathsf{Var}(X)}.$$

- Idea Use the samples $\{(X_k, Y_k), k \leq n\}$ to approximate $b_\star$ by $\hat{b}_n$.
- Warning : check that (asymptotically) the properties of $\hat{\mu}_n(\hat{b}_n)$ are the same than $\hat{\mu}_n(b_\star)$?
- Alternative (but costly) Use an independent sample $\{(X'_k, Y'_k), k \leq n_1\}$ to approximate $b_\star$.

Limiting distribution

- (show)

$$\frac{\sqrt{n}}{\sigma(b_\star)}\left(\hat{\mu}_n(\hat{b}_n) - \mathbb{E}[Y]\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

where

$$\hat{b}_n \stackrel{\text{def}}{=} \frac{\sum_{k=1}^n (X_k - \mathbb{E}[X])(Y_k - \bar{\mu}_n)}{\sum_{k=1}^n (X_k - \mathbb{E}[X])^2} \qquad b_\star \stackrel{\text{def}}{=} \frac{\mathsf{Cov}(X,Y)}{\mathsf{Var}(X)}$$

- Therefore $\hat{\mu}_n(\hat{b}_n)$ as the same asymptotic distribution $\hat{\mu}_n(b_\star)$.
- The asymptotic distribution is unchanged if $\sigma^2(b_\star)$ by a consistent estimator of this quantity.

Multidimensional extensions

- Several control values $X = (X^{(1)}, \cdots, X^{(d)})^T$ can be available
- In such case, the control variate estimate is

$$\hat{\mu}_n(b) \stackrel{\text{def}}{=} \bar{\mu}_n - b^T \left(\bar{X}_n - \mathbb{E}[X] \right)$$

where

$$\bar{X}_n \stackrel{\text{def}}{=} \frac{1}{n} \sum_{k=1}^{n} X_k \qquad\qquad X_k \text{ i.i.d. } \sim X$$

- Along the same lines the optimal weight is $b_\star \stackrel{\text{def}}{=} \Sigma_X^{-1} \Sigma_{X,Y}$ where

$$\text{Var}(\hat{\mu}_n(b_\star)) = \text{Var}(\bar{\mu}_n) \ \left(1 - R^2\right) \qquad R^2 \stackrel{\text{def}}{=} \frac{\Sigma_{X,Y}^T \Sigma_X^{-1} \Sigma_{X,Y}}{\text{Var}(Y)}$$

where Σ_X is the covariance matrix of X and $\Sigma_{X,Y}$ is the intercorrelation $\text{Cov}(X^{(k)}, Y)$.

Troisième partie III

Miscellaneous applications

Example 1

Objective

$$\mathcal{I} \stackrel{\text{def}}{=} \mathbb{E}\left[(W_1 + W_2)^{5/4}\right]$$

where W_1, W_2 are independent with the same Weibull distribution

$$f(x) = 3/2\sqrt{x}\exp(-x^{3/2})\mathbb{1}_{\mathbb{R}+}(x).$$

1. If U is uniform on $[0, 1]$, then $W = (-\ln U)^{2/3}$ is distributed as W_1.
2. Estimate $\mathcal{I}$ using a crude MC method.
3. Propose a control variate estimator of $\mathcal{I}$ using $X = U_1 U_2$ as a control variate.

Example 1

Objective

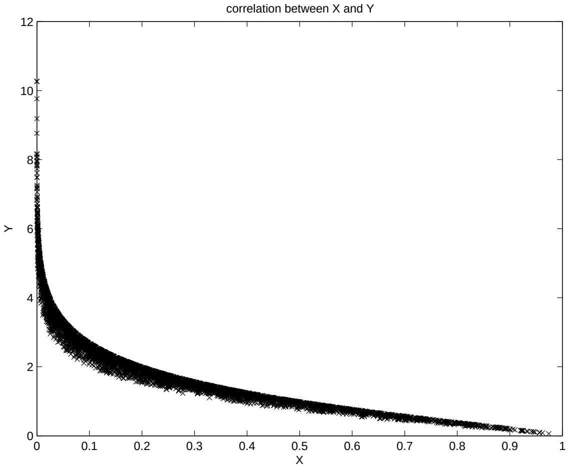
$$\mathcal{I} \overset{\text{def}}{=} \mathbb{E}\left[(W_1 + W_2)^{5/4}\right]$$

where W_1, W_2 are independent with the same Weibull distribution

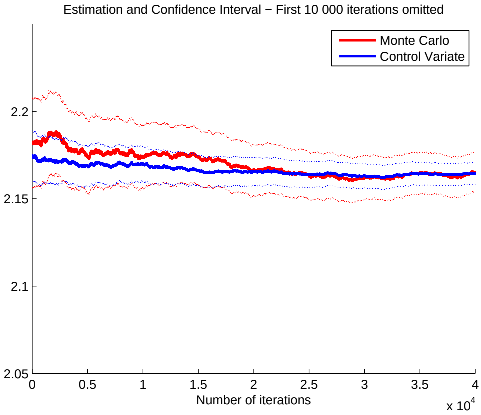
$$f(x) = 3/2\sqrt{x}\exp(-x^{3/2})\mathbb{1}_{\mathbb{R}+}(x).$$

1. If U is uniform on $[0,1]$, then $W = (-\ln U)^{2/3}$ is distributed as W_1.
2. Estimate $\mathcal{I}$ using a crude MC method.
3. Propose a control variate estimator of $\mathcal{I}$ using $X = U_1 U_2$ as a control variate. in this case
4. On estime : $\sqrt{1 - \rho^2_{X,Y}} \approx 0.55$.

Results


correlation between X and Y
Y
X
0
2
4
6
8
10
12
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

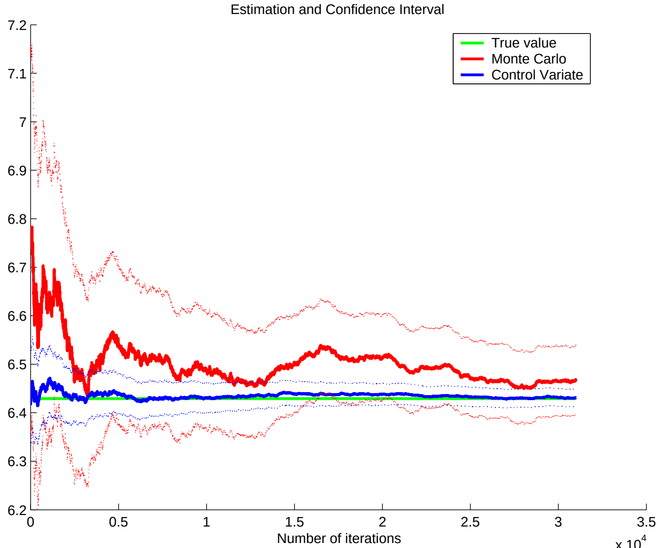
Results


Estimation and Confidence Interval − First 10 000 iterations omitted
Monte Carlo
Control Variate
2.2
2.15
2.1
2.05
0
0.5
1
1.5
2
2.5
3
3.5
4
Number of iterations
x 10^4

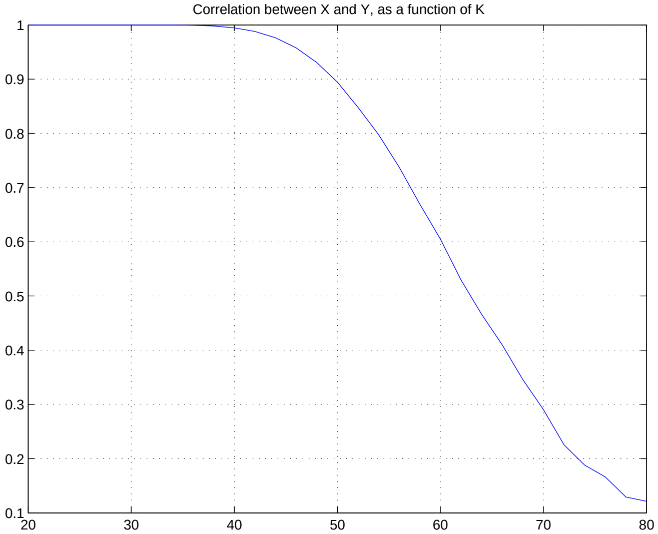
Call put-parity

- For all x, K
$$(x-K)_+ - (K-x)_+ = x - K$$
- When $\mathbb{E}[S]$ is known, a natural control variable for $\mathbb{E}[(S-K)_+]$ is $X = (S-K)$.
- In financial engineering
$$\mathcal{I} \overset{\text{def}}{=} \mathbb{E}\left[\exp(-rT)\ \ (S_T - K)_+\right]$$
where S_T is the price of the underlying asset at the maturity T, modeled has a GBM $S_0 = x$ where
$$\mathbb{E}\left[\exp(-rT)\ \ (S_T - K)\right] = x - \exp(-rT)K$$
- In this case
$$Y = \exp(-rT)\ \ (S_T - K)_+ \qquad\qquad X = \exp(-rT)\ \ (S_T - K)$$

Numerical applications



Numerical applications



Asian Option

$$\mathcal{I} \stackrel{\text{def}}{=} \exp(-rT)\, \mathbb{E}\left[\left(M^{-1}\sum_{k=1}^{M} S_{t_k} - K\right)_+\right]$$

where $t_k = kT/M$ and $\{S_t, t \geq 0\}$ is a GBM

$$S_t = S_0 \exp\left((r - 0.5\sigma^2)t + \sigma W_t\right).$$

Three possible control variates

$$X^{(1)} = \exp(-rT)S_T - S_0$$

$$X^{(2)} = \exp(-rT)\left(S_T - K\right)_+ - \mathbb{E}\left[\exp(-rT)\left(S_T - K\right)_+\right]$$

$$X^{(3)} = \exp(-rT)\left(\exp\left(M^{-1}\sum_{k=1}^{M} \ln S_{t_k}\right) - K\right)_+ - \mathbb{E}\left[\exp(-rT)\left(\exp\left(M^{-1}\sum_{k=1}^{M} \ln S_{t_k}\right) - K\right)_+\right].$$

Numerical Applications