Control variates with kernel smoothing : toward faster than root n rates

Benjamin Cohen, Arthur Gourrin, Julien Toucheboeuf, Hugues Rene-Bazin

Macs205a

Outline

Outline

Introduction

- Problem : estimating $I = \int g(x)f(x)dx$, with $f : \mathbb{R}^d \mapsto \mathbb{R}$ a density
- How : thanks to a Monte Carlo method
- Goal : get a faster convergence thanks to :
  - Control variates methods
  - Improvements with Kernel smoothing

Outline

Control variable estimator

- Estimator : $\hat{I}_n(\phi) = \frac{1}{n} \sum_{i=1}^{n} \frac{g(X_i)f(X_i) - \phi(X_i)}{q(X_i)}$
- with $\phi(x) = \sum_{k=1}^{m} \beta_k \phi_k(x)$ s.t. $\forall \int \phi_k = 0$.
- UnBiased estimator & Variance to minimise

research of an optimal $\hat{\phi}$

- Optimisation problem : $\hat{\beta} \in \underset{\beta \in \mathbb{R}^m}{\operatorname{argmin}} \frac{1}{n-1} \sum_{i=1}^{n} \left(\frac{f(X_i)g(X_i)}{q(X_i)} - \beta Z_i - \hat{I}_n(\beta) \right)^2$
- With $\forall i \in \{1,...,n\}, Z_i = (\frac{\phi_1(X_i)}{q(X_i)},..., \frac{\phi_m(X_i)}{q(X_i)})^T$.
- Hilbert projection Theorem : $\hat{\beta} = (Z_c^T Z_c)^{-1}(Z_c)G$.
- $Z_c = (Z_1 - \bar{Z}, Z_2 - \bar{Z}, ..., Z_n - \bar{Z})^T$, $\bar{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i$
- $G = (\frac{f(X_1)g(X_1)}{q(X_1)},..., \frac{f(X_n)g(X_n)}{q(X_n)})^T$

Asymptotic behavior

Asymptotic convergence

$$\sqrt{n}(\hat{I}_n^{cv}(\hat{\beta}) - I) \xrightarrow{d} \mathcal{N}(0, \sigma_m^2)$$

So :

- $\sigma_m^2 = \min_{\beta \in \mathbb{R}^m} V(g(X_1)f(X_1) - \beta^T Z_1)$
- Result due to the TCL, and the Slutsky lemma
- Z as it was defined in the previous slide

Numeric Simulation

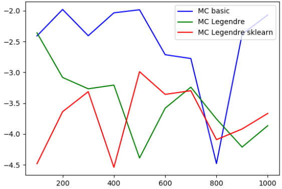

Figure: Plot of the log error for the three methods (1000 points)

- MC_Legendre : OLS with the theoretical formula
- sklearn With the linear_regression method

Outline

General Kernel smoothing method

- Kernels : for $h > 0, \int K_h = 1$
- General idea, approximate a function u, by $\hat{u}(x) = \frac{1}{n}\sum_i K_h(x - X_i)$
- Problem : $\mathbb{E}[\hat{u}] = 1$. It means that it is not a control variate
- Control variate : $h_x(X) = \frac{K_h(x-X)}{q(X)} - 1$.

New Optimisation problem

- $\hat{\alpha} \in \underset{\alpha \in \mathbb{R}^m}{\operatorname{argmin}} \sum_i \left(\frac{f(X_i)g(X_i)}{q(X_i)} - \sum_{j=1}^{m} \alpha_j \left(\frac{K_h(X_i - X_j)}{q(X_i)} - 1 \right) \right)^2$

Numeric simulation

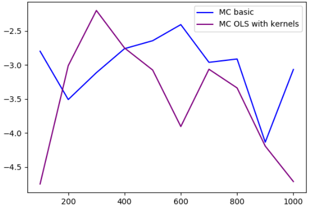

Figure: Kernel Smoothing vs MC Naif (dim 1)

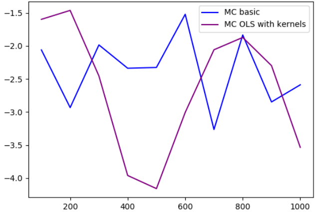

Figure: Kernel Smoothing vs MC Naif (dim 10)

Comparison Legendre versus Kernel approach

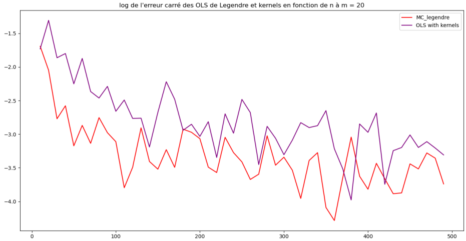

Figure: Log square error of Legendre & Kernels

Outline

Estimator $\tilde{\phi}_m(x)$

- New estimator of gf: $\tilde{\phi}_m(x) = \frac{1}{m}\sum_{i=1}^{m} \frac{g(X_j')f(X_j')}{q(X_j')} K_h(x - X_j')$
- Mathematical model : Based on Kernels
- Solution : no OLS/LASSO this time

Bias & Variance of $\tilde{\phi}_m(x)$

Bias

$$\text{bias} = gf(x) + \int (gf(x-hu) - gf(x)) \times K(u)du$$

- $\int (gf(x-hu) - gf(x))K(u)du \leq Lh\mathbb{E}[K_h] \times 1 \xrightarrow{h\to 0} 0$
- unbiased estimator when $h \to 0$.

Variance

$$\sigma_K^2(x) = \frac{1}{m} V_q\left(\frac{g(X_1')f(X_1')}{q(X_1')} K_h(x - X_1')\right)$$

From $\tilde{\phi}_m(x)$ to a control variate

We have almost an unbiased estimator of gf, but not yet a control variate :

$$\mathbb{E}[\frac{\tilde{\phi}_m}{q}(X)] = \frac{1}{m}\sum_{j=1}^{m}\frac{g(X_j')f(X_j')}{q(X_j')} \times 1 = \hat{\mu}_m$$

Solution:

- Take : $\forall i \in \{1, ..., n\}, Z_{m,i} = \frac{\tilde{\phi}_m}{q}(X_i) - \hat{\mu}_m$
- Estimator :

$$\hat{I}_n^{cv}(\tilde{\phi}_m) = \frac{1}{n}\sum_{i=1}^{n}\frac{g(X_i)f(X_i) - Z_{m,i}}{q(X_i)}$$

Mean & Variance of $\hat{I}_n^{cv}(\tilde{\phi}_m)$

Mean

$$\mathbb{E}[\hat{I}_n^{cv}(\tilde{\phi}_m)] = \mathbb{E}\left[\frac{g(X_1)f(X_1) - Z_{m,1}}{q(X_1)}\right] = I$$

Variance

$$V(\hat{I}_n^{cv}(\tilde{\phi}_m)) = \frac{V(\frac{g(X_1)f(X_1)}{q(X_1)}) + V(\frac{\phi(X_1)}{q(X_1)})}{n} - 2Cov\left(\frac{g(X_1)f(X_1)}{q(X_1)}, \frac{\phi(X_1)}{q(X_1)^2}\right) - I^2$$

Outline

Simulation method

Two method were implemented :

- the OLS
- LASSO method

Numeric estimation of Monte Carlo I

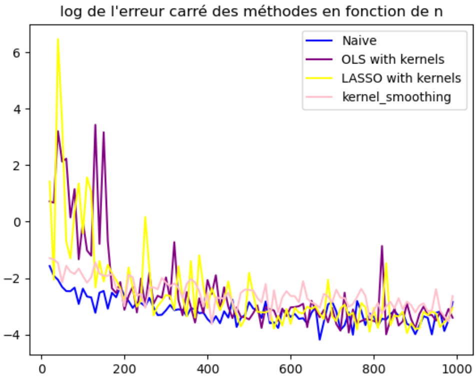

Figure: Log square error

Outline

Conclusions and future works

- 3 different theoretical methods used
- Simulations made relying of thoose methods
- results : not really concluding