The data-driven COS method

Application to option pricing

A.Leitao, L. Ortiz-Gracia, C. W. Oosterlee and S. M. Bohte

Delft University of Technology - Centrum Wiskunde & Informatica

UB - December 2, 2016

Motivation

- Combine the ideas behind the COS method and Monte Carlo simulation.
- Preserving the individual advantages and overcoming the particular disadvantages.
- Make the COS method more generally and directly applicable and more flexible.
- Improve the convergence of the Monte Carlo method.

Definitions

Option

A contract that offers the buyer the right, but not the obligation, to buy (call) or sell (put) a financial asset at an agreed-upon price (the strike price) during a certain period of time or on a specific date (exercise date). Investopedia.

Option price

The fair value to enter in the option contract. In other words, the (discounted) expected value of the contract.

$$V_\tau = D_\tau \mathbb{E}\left[v(S(\tau))\right]$$

where $v(\cdot)$ is the *payoff* function, $S(\tau)$ the future value of an underlying asset, $S(t)$, and D_τ the discount factor.

Definitions - cont.

Pricing techniques

- Stochastic process, $S(t)$: Stochastic differential equation (SDE).
- Simulation: Monte Carlo method.
- Fourier-based methods.
- PDEs: Feynman-Kac theorem.

Types of options - Exercise time

- European: End of the contract, $\tau = T$.
- Early-exercise: American($\tau \in [t, T]$) or Bermudan($\tau \in \{t1, \ldots, t_M\}$).
- Many others: Asian, barrier, ...

Types of options - Payoff

- Vanilla: $[c(S(\tau) - K)]^+$, call($c = 1$) or put($c = -1$).
- Many others: Digital, Gap, ...

The COS method

- A lot of work behind: [FO08], [FO09], etc.
- Fourier-based method to price options.
- Point of departure is risk-neutral valuation formula:

$$v(x,t) = \mathrm{e}^{-r(\tau-t)}\mathbb{E}\left[v(y,\tau)|x\right] = \mathrm{e}^{-r(\tau-t)}\int_{\mathbb{R}} v(y,\tau)f(y|x)\mathrm{d}y,$$

  where r is the risk-free rate and $f(y|x)$ is the density of the underlying process. Typically, we have:

$$x := \log\left(\frac{S(t)}{K}\right) \quad \text{and} \quad y := \log\left(\frac{S(t+\tau)}{K}\right).$$

- $f(y|x)$ is unknown in most of cases.
- However, characteristic function available for many models.
- Exploit the relation between the density and the characteristic function (Fourier pair).

Pricing European options with the COS method

- For a function supported on $[0, \pi]$, the cosine series expansion reads

$$f(\theta) = \sum_{k=0}^{\infty}{}' A_k \cdot \cos(k\theta).$$

where $\sum'$ indicates that the first term is weighted by one-half.

- Other finite interval $[a, b]$, change of variables:

$$\theta := \pi \frac{x-a}{b-a}.$$

- How to compute the support $[a, b]$ for a particular problem is crucial.
- The COS method relies on a *cumulant* approach.

Pricing European options with the COS method

- The cosine series expansion of $f(y|x)$ in the support $[a, b]$ is

$$f(y|x) = {\sum_{k=0}^{\infty}}' A_k(x) \cdot \cos\left(k\pi \frac{y-a}{b-a}\right).$$

- The option value, $v(x, t)$ with $\tau = T$, can be then approximated by

$$v(x,t) \approx \mathrm{e}^{-r(T-t)} \int_a^b v(y, T) {\sum_{k=0}^{\infty}}' A_k(x) \cos\left(k\pi \frac{y-a}{b-a}\right) \mathrm{d}y.$$

- Interchanging sum and integration, and introducing the definition

$$V_k := \frac{2}{b-a} \int_a^b v(y, T) \cos\left(k\pi \frac{y-a}{b-a}\right) \mathrm{d}y,$$

an approximated pricing formula is obtained (after series truncation)

$$v(x,t) \approx \frac{1}{2}(b-a)\mathrm{e}^{-r(T-t)} {\sum_{k=0}^{N-1}}' A_k(x) V_k.$$

Pricing European options with the COS method

- The $A_k(x)$ expansion coefficients are

$$A_k(x) = \frac{2}{b-a} \int_a^b f(y|x) \cos\left(k\pi \frac{y-a}{b-a}\right) \mathrm{d}y.$$

- By employing the Fourier transform properties and based on the characteristic function, $\phi(u;x)$, associated to $f(y|x)$:

$$A_k(x) \approx \frac{2}{b-a} \mathcal{R}\left\{\phi\left(\frac{k\pi}{b-a};x\right) \cdot \exp\left(-i\frac{k\pi a}{b-a}\right)\right\}.$$

- The COS pricing formula for European options

$$v(x,t) \approx \mathrm{e}^{-r(T-t)} {\sum_{k=0}^{N-1}}{}' \mathcal{R}\left\{\phi\left(\frac{k\pi}{b-a};x\right) \cdot \exp\left(-i\frac{k\pi a}{b-a}\right)\right\} V_k.$$

- The V_k coefficients are known for many types of payoffs.

Pricing Bermudan options with the COS method

- A Bermudan option can be exercised at a set of predefined dates.
- The price is computed by using the risk-neutral valuation formula.
- With M exercise dates $t_1 < \cdots < t_M = T$ and with $\Delta t = t_m - t_{m-1}$, the pricing formula for Bermudan option then reads

$$
\begin{aligned}
c(x, t_{m-1}) &= \mathrm{e}^{-\Delta t} \int_{\mathbb{R}} v(y, t_m) f(y|x) \mathrm{d}y, \\
v(x, t_{m-1}) &= \max\left(g(x, t_{m-1}), c(x, t_{m-1})\right),
\end{aligned}
$$

applied recursively, starting in $t_m = T$ until t_2, and followed by

$$
v(x, t_0) = \mathrm{e}^{-\Delta t} \int_{\mathbb{R}} v(y, t_1) f(y|x) \mathrm{d}y.
$$

- The functions $v(x, t)$, $c(x, t)$ and $g(x, t)$ are the option value, the continuation value and the payoff value at time t, respectively.

Pricing Bermudan options with the COS method

- Following a similar procedure as in the case of the European options, the continuation value and the option value can be approximated as

$$c(x, t_{m-1}) \approx \mathrm{e}^{-\Delta t} \sum_{k=0}^{N-1}{}' \mathcal{R}\left\{\phi\left(\frac{k\pi}{b-a}; x\right) \cdot \exp\left(-i\frac{k\pi a}{b-a}\right)\right\} V_k(t_m),$$

and

$$v(x, t_0) \approx \mathrm{e}^{-\Delta t} \sum_{k=0}^{N-1}{}' \mathcal{R}\left\{\phi\left(\frac{k\pi}{b-a}; x\right) \cdot \exp\left(-i\frac{k\pi a}{b-a}\right)\right\} V_k(t_1),$$

where, again, $\phi(u; x)$ is the characteristic function.

- Pricing a Bermudan option is reduced to the computation of $V_k(t_1)$.
- The coefficients V_k at any time t_m can be defined by

$$V_k(t_m) := \frac{2}{b-a} \int_a^b v(y, t_m) \cos\left(k\pi \frac{y-a}{b-a}\right) \mathrm{d}y.$$

"Learning" densities from data

- *Statistical learning theory*: deals with the problem of finding a predictive function based on data. Wikipedia.
- We follow the analysis about the problem of density estimation proposed by Vapnik in [Vap98].
- Given independent and identical distributed samples $X_1, X_2, \ldots, X_n$.
- By definition, density $f(x)$ is related to the *cumulative distribution function*, $F(x)$, by means of the expression

$$\int_{-\infty}^{x} f(y)\mathrm{d}y = F(x).$$

- $F(x)$ can be very accurately approximated by the empirical equivalent

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n}\eta(x - X_i),$$

where $\eta(\cdot)$ is the step-function. Convergence $O(1/\sqrt{n})$.

Regularization approach

- The previous equation can be rewritten as a linear operator equation

$$Af = F \approx F_n,$$

where the operator $Az = \int_{-\infty}^{x} z\mathrm{d}z$.
- Stochastic ill-posed problem. *Regularization method* (Vapnik).
- Given a lower semi-continuous functional $W(f)$ such that:
  - Solution of $Af = F_n$ belongs to $\mathcal{D}$, the domain of definition of $W(f)$.
  - The functional $W(f)$ takes real non-negative values in $\mathcal{D}$.
  - The set $\mathcal{M}_c = \{f : W(f) \leq c\}$ is compact in $\mathcal{H}$ (the space where the solution exits and is unique).
- Then we can construct the functional

$$R_{\gamma_n}(f, F_n) = L^2_{\mathcal{H}}(Af, F_n) + \gamma_n W(f),$$

where $L_{\mathcal{H}}$ is a metric of the space $\mathcal{H}$ (loss function) and γ_n is the parameter of regularization satisfying that $\gamma_n \to 0$ as $n \to \infty$.
- Under these conditions, a function f_n minimizing the functional converges almost surely to the desired one.

The regularization approach and the COS method

- Assuming $f(x)$ belongs to the functions whose p-th derivatives belong to $L_2(0,\pi)$. We consider the risk functional as the form

$$R_{\gamma_n}(f, F_n) = \int_0^{\pi} \left(\int_0^{x} f(y)\mathrm{d}y - F_n(x) \right)^2 \mathrm{d}x + \gamma_n \int_0^{\pi} \left(f^{(p)}(x) \right)^2 \mathrm{d}x.$$

- Assuming the solution is in the form (as in the COS method)

$$f_n(\theta) = {\sum_{k=0}^{\infty}}' \hat{A}_k \cos(k\theta),$$

where $\hat{A}_0, \hat{A}_1, \ldots, \hat{A}_{k-1}, \ldots$ are the expansion coefficients.

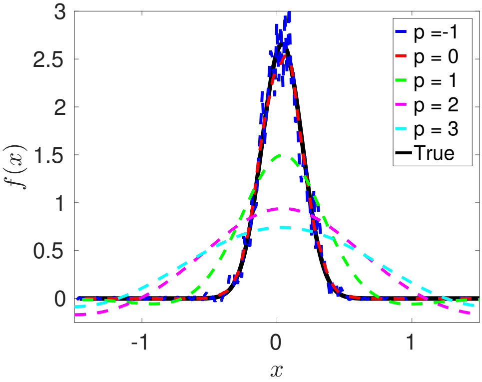
- Plugging the expansion in the risk functional, it can be proved that the minimum of $R_{\gamma_n}(f_n, F_n)$ is reached when

$$\hat{A}_k = \frac{2}{\pi} \cdot \frac{\frac{1}{n}\sum_{i=1}^{n} \cos(k\theta_i)}{1 + \gamma_n k^{2(p+1)}}, k = 0, 1, 2, \ldots.$$

where, again, n is the number of available samples.

Choice of the free parameters: γ_n and p

- The choice of optimal values of γ_n and p is crucial in terms of accuracy and efficiency.
- There is no rule or procedure to obtain an optimal p.
- As a rule of thumb, $p = 0$ seems to be the most appropriate value.
- Fixing p, we rely on the computation of an optimal γ_n.

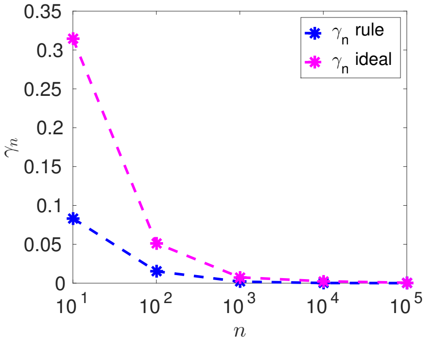


Choice of γ_n

- For the regularization parameter γ_n, a rule that ensures asymptotic convergence

$$\gamma_n = \frac{\log \log n}{n}.$$

- But it is not the optimal value of γ_n, i.e. the one which provides the fastest convergence in practical situations.



Choice of γ_n

- Exploit the relation between the empirical and real (unknown) CDFs.
- This relation can be modeled by *statistical laws* or *statistics*: Kolmogorov-Smirnov, Anderson-Darling, Smirnov-Cramér–von Mises.
- Preferable: a measure of the distance between the $F_n(x)$ and $F(x)$ follows a known distribution.
- We have chosen Smirnov-Cramér–von Mises(SCvM):

$$\omega^2 = n \int_{\mathbb{R}} (F(x) - F_n(x))^2 \, \mathrm{d}F(x).$$

- Assume we have an approximation, F_{γ_n} (which depends on γ_n).
- An *almost* optimal γ_n is computed by solving the equation

$$\sum_{i=1}^{n} \left(F_{\gamma_n}(\bar{X}_i) - \frac{i - 0.5}{n} \right)^2 = m_S - \frac{i}{12n},$$

where $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_n$ is the ordered array of samples $X_1, X_2, \ldots, X_n$ and m_S the mean of the ω^2.

Influence of γ_n

- To assess the impact of γ_n: *Mean integrated Squared Error* (MiSE):

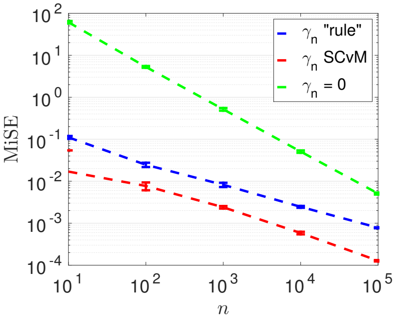
$$\mathbb{E}\left[\|f_n - f\|_2^2\right] = \mathbb{E}\left[\int_{\mathbb{R}} (f_n(x) - f(x))^2 \,\mathrm{d}x\right].$$

- A formula for the MiSE formula is derived in our context:

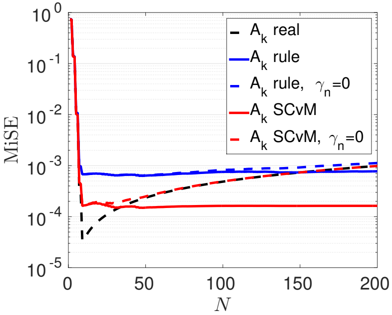
$$\mathrm{MiSE} = \frac{1}{n}\sum_{k=1}^{N} \frac{1}{\left(1+\gamma_n k^{2(p+1)}\right)^2}\left(\frac{1}{2} + \frac{1}{2}A_{2k} - A_k^2\right) + \sum_{k=N+1}^{\infty} A_k^2.$$

- Two main aspects influenced γ_n: accuracy in n and stability in N.

Influence of γ_n



(a) Convergence in terms of n



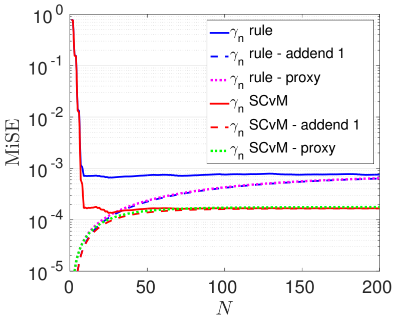
(b) Convergence in terms of N

Figure: Influence of γ_n: .

Optimal number of terms N

- Try to find a *minimum optimal* value of N.
- N considerably affects the performance.
- We wish to avoid the computation of any $\hat{A}_k$.
- We define a proxy for the MiSE and follow:

$$\text{MiSE} \approx \frac{1}{n} \sum_{k=1}^{N} \frac{\frac{1}{2}}{\left(1 + \gamma_n k^{2(p+1)}\right)^2}.$$



Optimal number of terms N

Data: n, γ_n

$N_{min} = 5$

$N_{max} = \infty$

$\epsilon = \frac{1}{\sqrt{n}}$

$\text{MiSE}_{prev} = \infty$

for $N = N_{min} : N_{max}$ **do**

$\quad \text{MiSE}_N = \frac{1}{n} \sum_{k=1}^{N} \frac{\frac{1}{2}}{\left(1+\gamma_n k^{2(p+1)}\right)^2}$

$\quad \epsilon_N = \frac{|\text{MiSE}_N - \text{MiSE}_{prev}|}{|\text{MiSE}_N|}$

$\quad$ **if** $\epsilon_N > \epsilon$ **then**

$\quad\quad N_{op} = N$

$\quad$ **else**

$\quad\quad$ Break

$\quad \text{MiSE}_{prev} = \text{MiSE}_N$

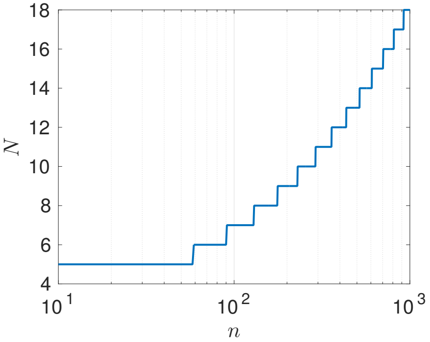


Figure: Almost optimal N.

The data-driven COS method

- Employ the regularization approach for PDF estimation in the COS framework.
- In both, the PDF is assumed to be in the form of a cosine series expansion.
- The minimum of the functional is in terms of the expansion coefficients.
- Take advantage of the COS machinery: pricing options.
- The samples are generated by Monte Carlo (risk-neutral measure).

The data-driven COS method - European options

- Key idea: the data-based $\hat{A}_k$ approximates the $A_k(x)$ in COS method.
- Let assume that we have the samples $S_i(t)$.
- As in COS method, a logarithmic transformation is made

$$Y_j := \log\left(\frac{S_j(T)}{K}\right).$$

- Due to the solution is defined in $(0, \pi)$, we further transform the samples as

$$\theta_i := \pi\frac{Y_i - a}{b - a}.$$

where the quantities a and b are defined as

$$a := \min_{1\leq j\leq n}(Y_j), \quad b := \max_{1\leq j\leq n}(Y_j).$$

- The samples, Y_i, must intrinsically consider the dependency on x of the function-like coefficients, $A_k(x)$.
- This is fulfilled when generated by the Monte Carlo method.

The data-driven COS method - European options

- New expression for the data-driven coefficients, $\hat{A}_k$:

$$\hat{A}_k = \frac{2}{b-a} \cdot \frac{\frac{1}{n}\sum_{i=1}^{n} \cos\left(k\pi \frac{Y_i - a}{b-a}\right)}{1 + \gamma_n k^{2(p+1)}}, k = 1, 2, \ldots .$$

- By substituting the $A_k(x)$ in the COS formula by the $\hat{A}_k$ coefficients, we obtain the ddCOS pricing formula for European options

$$\begin{aligned}\hat{v}(x,t) &= \frac{1}{2}(b-a)\mathrm{e}^{-r(T-t)} {\sum_{k=0}^{N-1}}' \hat{A}_k V_k, \\ &= \mathrm{e}^{-r(T-t)} {\sum_{k=0}^{N-1}}' \frac{\frac{1}{n}\sum_{i=1}^{n} \cos\left(k\pi \frac{Y_i - a}{b-a}\right)}{1 + \gamma_n k^{2(p+1)}} \cdot V_k.\end{aligned}$$

The ddCOS method vs. Monte Carlo

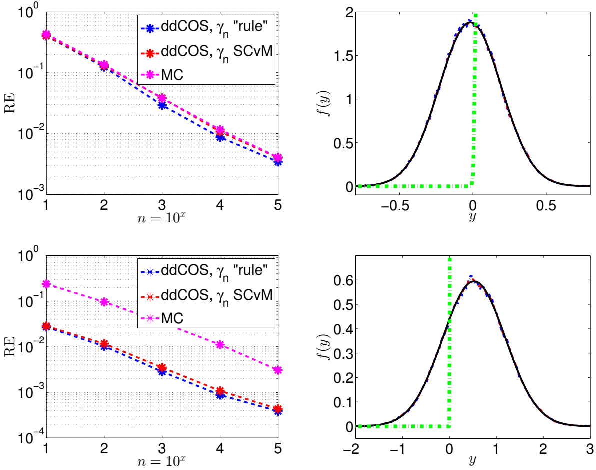


Figure: Convergence GBM.

The ddCOS method vs. Monte Carlo

- With more involved models: Jump-diffusion Merton model.

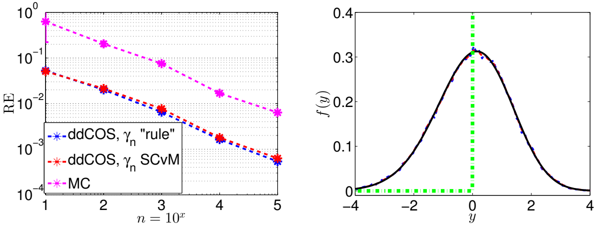


Figure: Convergence Merton.

The ddCOS method vs. Monte Carlo

- Almost optimal γ_n (SCvM) does not provide better results.
- Natural question: is the ddCOS method worth to use?
- Other words: is it better in terms of computational cost?

RE	$< 10^{-1}$	$< 10^{-2}$	$< 10^{-3}$
		GBM	
MC	$0.0095(10^3)$	$0.0147(10^5)$	$0.6721(10^7)$
ddCOS	$0.0256(10^1)$	$0.0258(10^3)$	$0.2696(10^5)$
Speedup	$\times 0.37$	$\times 0.57$	$\times 2.49$
		Merton	
MC	$0.0396(10^3)$	$0.1315(10^5)$	$13.6055(10^7)$
ddCOS	$0.0527(10^1)$	$0.0558(10^3)$	$0.3861(10^5)$
Speedup	$\times 0.75$	$\times 2.36$	$\times 35.24$

Table: Time(s) vs. accuracy. In parentheses, the number of required samples, n.

The ddCOS method - Pricing European options

K(% of S_0)	80%	90%	100%	110%	120%
BS	24.0784	15.4672	8.5917	4.0759	1.6600
ddCOS	24.0799	15.4730	8.6042	4.0827	1.6668
MSE	5.7344×10^{-5}				

Table: GBM. European call, $S_0 = 100$, $T = 1$, $r = 0.05$, $\sigma = 0.15$.

K(% of S_0)	80%	90%	100%	110%	120%
Merton	62.7649	59.5454	56.5534	53.7676	51.1694
ddCOS	62.7962	59.5713	56.5689	53.7736	51.1659
MSE	3.8782×10^{-4}				

Table: Merton. European call, $S_0 = 100$, $T = 5$, $r = 0.1$, $\sigma = 0.3$, $\lambda = 3$, $\mu_J = -0.2$, $\sigma_J = 0.2$.

The ddCOS method - Pricing European options

$K(\%$ of $S_0)$	80%	90%	100%	110%	120%
MC	34.5093	26.3220	18.1347	9.9550	2.6402
ddCOS	34.5015	26.3142	18.1269	9.9473	2.6337
MSE	2.3440×10^{-6}				

Table: CEV. European call, $S_0 = 100$, $T = 2$, $r = 0.1$, $\sigma = 0.3$, $\beta = 0.5$.

$K(\%$ of $S_0)$	80%	90%	100%	110%	120%
MC	23.9017	14.4483	5.6862	0.6043	0.0019
ddCOS	23.9043	14.4510	5.6900	0.6092	0.0043
MSE	1.1685×10^{-5}				

Table: SABR. European call, $S_0 = 100$, $T = 1$, $r = 0.05$, $\sigma_0 = 0.15$, $\alpha = 0.3$, $\beta = 0.8$, $\rho = -0.8$.

The data-driven COS method - Bermudan options

- Pricing Bermudan options under the ddCOS is more involved: several exercise times, t_m.
- The characteristic function now appears in the computation of both, the continuation value a the final option value.
- The "conditionalty" in the samples is not straightforward since the current state at t_m is conditional on the previous one at t_{m-1} (not on the initial state S_0).
- For pricing Bermudan option, the state variables x and y are defined

$$x := \log\left(\frac{S(t_{m-1})}{K}\right) \quad \text{and} \quad y := \log\left(\frac{S(t_m)}{K}\right).$$

The data-driven COS method - Bermudan options

- We propose the following approximation

$$
\begin{aligned}
f(y|x) &\overset{\mathrm{d}}{\approx} f(y-x) \\
&\overset{\mathrm{d}}{=} f\left(\log\left(\frac{S(t_m)}{K}\right) - \log\left(\frac{S(t_{m-1})}{K}\right)\right) \\
&\overset{\mathrm{d}}{=} f\left(\log\left(\frac{S(t_m)}{S(t_{m-1})}\right)\right) \overset{\mathrm{d}}{=} f\left(\log\left(\frac{S(t_1)}{S(t_0)}\right)\right).
\end{aligned}
$$

  where $\overset{\mathrm{d}}{=}$ indicates equality in the distribution sense.
- If we consider a particular realization, x', then

$$
f(y|x = \mathrm{x}') \overset{\mathrm{d}}{\approx} \mathrm{f}\left(\log\left(\frac{\mathrm{S}(\mathrm{t}_1)}{\mathrm{S}(\mathrm{t}_0)}\right)\right) + \mathrm{x}'.
$$

The data-driven COS method - Bermudan options

- Schematic representation of the idea:

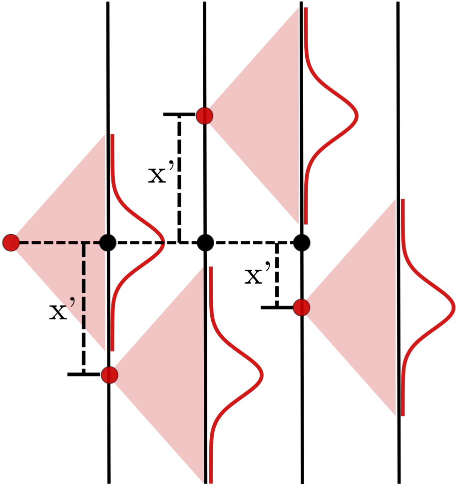


Figure: Approximation to the continuation value computation.

The data-driven COS method - Bermudan options

- According to that, we define the samples as

$$Z_j = \log\left(\frac{S_j(t_1)}{S(t_0)}\right),$$

- By applying the ddCOS method we have

$$\hat{A}_k(x) = \frac{\frac{1}{n}\sum_{j=1}^{n}\cos\left(k\pi\frac{(Z_j+x)-a}{b-a}\right)}{1+\gamma_n k^{2(p+1)}}.$$

- Then, the data-based expression for the continuation value

$$\begin{aligned}\hat{c}(x,t_{m-1}) &= \exp(-\Delta t)\sum_{k=0}^{N-1}{}' \hat{B}_k(x)V_k(t_m) \\ &= \exp(-\Delta t)\sum_{k=0}^{N-1}{}' \frac{\frac{1}{n}\sum_{i=1}^{n}\cos\left(k\pi\frac{(Z_i(t_1)+x)-a}{b-a}\right)}{1+\gamma_n k^{2(p+1)}}\cdot V_k(t_m).\end{aligned}$$

The ddCOS method - Pricing Bermudan options

K(% of S_0)	80%	90%	100%	110%	120%
COS	1.6413	3.8766	7.6122	13.0919	20.3759
ddCOS	1.6400	3.8721	7.6011	13.0756	20.3730
MSE	8.3930×10^{-5}				

Table: GBM. Bermudan put, $S_0 = 100$, $T = 2$, $r = 0.05$, $\sigma = 0.2$.

K(% of S_0)	80%	90%	100%	110%	120%
COS	13.2743	17.8149	22.9726	28.7044	34.9664
ddCOS	13.1908	17.7128	22.8526	28.5668	34.8134
MSE	1.4828×10^{-2}				

Table: Merton. Bermudan put, $S_0 = 100$, $T = 2$, $r = 0.05$, $\sigma = 0.2$, $\lambda = 3$, $\mu_J = -0.2$, $\sigma_J = 0.2$.

Conclusions

- We have combined a density estimation procedure with the COS method.
- Resulting in a simple and very efficient technique for option pricing.
- The ddCOS method improves the convergence w.r.t. Monte Carlo.
- For high accuracy, faster than Monte Carlo.
- Ongoing work:
  - Case where the use of γ_n by SCvM is justified.
  - "Variance reduction" for the ddCOS method.
  - Bermudan approximation.
  - 2D extension.

References

Fang Fang and Cornelis W. Oosterlee.
A novel pricing method for European options based on Fourier-cosine series expansions.
SIAM Journal on Scientific Computing, 31:826–848, November 2008.

Fang Fang and Cornelis W. Oosterlee.
Pricing early-exercise and discrete Barrier options by Fourier-cosine series expansions.
Numerische Mathematik, 114(1):27–62, 2009.

Vladimir N. Vapnik.
Statistical Learning Theory.
Wiley-Interscience, 1998.

Suggestions, comments & questions

Thank you for your attention