The data-driven COS method

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

Delft University of Technology - Centrum Wiskunde & Informatica

CMMSE 2017, July 6, 2017
Outline

1. The COS method
2. “Learning” densities
3. The data-driven COS (ddCOS) method
4. Applications of the ddCOS method
5. Conclusions
The COS method

- Well known and established method: [FO08], [FO09], etc.
- Fourier-based method to price financial options.
- Starting point is risk-neutral valuation formula:

\[ v(x, t) = e^{-r(T-t)} \mathbb{E}[v(y, T)|x] = e^{-r(T-t)} \int_{\mathbb{R}} v(y, T)f(y|x)dy, \]

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

\[ x := \log \left( \frac{S(0)}{K} \right) \quad \text{and} \quad y := \log \left( \frac{S(T)}{K} \right), \]

- \( f(y|x) \) 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).
The COS method - European options

- $f(y|x)$ is approximated, on a finite interval $[a, b]$, by a cosine series
  \[
  f(y|x) = \frac{1}{b-a} \left( A_0 + 2 \sum_{k=1}^{\infty} A_k(x) \cdot \cos \left( k\pi \frac{y-a}{b-a} \right) \right),
  \]
  
  $A_0 = 1$, $A_k(x) = \int_a^b f(y|x) \cos \left( k\pi \frac{y-a}{b-a} \right) dy$, $k = 1, 2, \ldots$

- Interchanging the summation and integration and introducing the definition
  \[
  V_k := 2 \int_a^b \nu(y, T) \cos \left( k\pi \frac{y-a}{b-a} \right) dy,
  \]
  we find that the option value is given by
  \[
  \nu(x, t) \approx e^{-r(T-t)} \sum_{k=0}^{\infty} A_k(x) V_k,
  \]
  where ′ indicates that the first term is divided by two.
Pricing European options with the COS method

- Coefficients $A_k$ can be computed from the ChF.
- Coefficients $V_k$ are known analytically (for many types of options).
- Closed-form expressions for the option Greeks $\Delta$ and $\Gamma$

$$\Delta = \frac{\partial v(x, t)}{\partial S} = \frac{1}{S(0)} \frac{\partial v(x, t)}{\partial x} \approx \exp(-r(T-t)) \sum_{k=0}^{\infty} \frac{\partial A_k(x)}{\partial x} \frac{V_k}{S(0)},$$

$$\Gamma = \frac{\partial^2 v(x, t)}{\partial S^2} = \approx \exp(-r(T-t)) \sum_{k=0}^{\infty} \left(- \frac{\partial A_k(x)}{\partial x} + \frac{\partial^2 A_k(x)}{\partial x^2}\right) \frac{V_k}{S^2(0)}$$

- Due to the rapid decay of the coefficients, $v(x, t)$, $\Delta$ and $\Gamma$ can be approximated with high accuracy by truncating to $N$ terms.
“Learning” densities

- **Statistical learning theory**: deals with the problem of finding a predictive function based on data.
- We follow the analysis about the problem of density estimation proposed by Vapnik in [Vap98].
- Given independent and identically 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

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

- Function $F(x)$ is approximated by the empirical approximation

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

  where $\eta(\cdot)$ is the step-function. Convergence $O(1/\sqrt{n})$. 
The previous equation can be rewritten as a linear operator equation

\[ Cf = F \approx F_n, \]

where the operator \( Ch := \int_{-\infty}^{x} h(z)dz. \)


Given a lower semi-continuous functional \( W(f) \) such that:

- Solution of \( Cf = F_n \) belongs to \( D \), the domain of definition of \( W(f) \).
- The functional \( W(f) \) takes real non-negative values in \( D \).
- The set \( M_c = \{ f : W(f) \leq c \} \) is compact in \( \mathcal{H} \) (the space where the solution exits and is unique).

Then we can construct the functional

\[ R_{\gamma_n}(f, F_n) = L_{\mathcal{H}}^2(Cf, F_n) + \gamma_n W(f), \]

where \( L_{\mathcal{H}} \) is a metric of the space \( \mathcal{H} \) (loss function) and \( \gamma_n \) 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.
Assume $f(x)$ belongs to the functions whose $p$-th derivatives belong to $L_2(0, \pi)$, the kernel $\mathcal{K}(z - x)$ and

$$\mathcal{W}(f) = \int_{\mathbb{R}} \left( \int_{\mathbb{R}} \mathcal{K}(z - x)f(x)dx \right)^2 dz,$$

The risk functional

$$R_{\gamma_n}(f, F_n) = \int_{\mathbb{R}} \left( \int_0^x f(y)dy - F_n(x) \right)^2 dx + \gamma_n \int_{\mathbb{R}} \left( \int_{\mathbb{R}} \mathcal{K}(z - x)f(x)dx \right)^2 dz.$$

Denoting by $\hat{f}(u)$, $\hat{F}_n(u)$ and $\hat{\mathcal{K}}(u)$ the Fourier transforms, by definition

$$\hat{F}_n(u) = \frac{1}{2\pi} \int_{\mathbb{R}} F_n(x)e^{-iux}dx$$

$$= \frac{1}{2n\pi} \int_{\mathbb{R}} \sum_{j=1}^n \eta(x - X_j)e^{-iux}dx = \frac{1}{n} \sum_{j=1}^n \exp\left(-\frac{iuX_j}{iu}\right),$$

where $i = \sqrt{-1}$ is the imaginary unit.
Regularization and Fourier-based density estimators

- By employing the convolution theorem and Parseval’s identity

\[ R_{\gamma_n}(f, F_n) = \left\| \frac{\hat{f}(u) - \frac{1}{n} \sum_{j=1}^{n} \exp(-iuX_j)}{iu} \right\|_{L_2}^2 + \gamma_n \left\| \hat{K}(u)\hat{f}(u) \right\|_{L_2}^2. \]

- The condition to minimize \( R_{\gamma_n}(f, F_n) \) is given by,

\[
\frac{\hat{f}(u)}{u^2} - \frac{1}{nu^2} \sum_{j=1}^{n} \exp(-iuX_j) + \gamma_n \hat{K}(u)\hat{K}(-u)\hat{f}(u) = 0,
\]

which gives us,

\[
\hat{f}_n(u) = \left( \frac{1}{1 + \gamma_n u^2 \hat{K}(u)\hat{K}(-u)} \right) \frac{1}{n} \sum_{j=1}^{n} \exp(-iuX_j).
\]
\[ K(x) = \delta^{(p)}(x), \text{ and the desired PDF, } f(x) \text{ and its } p \text{-th derivative } (p \geq 0) \text{ belongs to } L_2(0, \pi), \text{ the risk functional becomes} \]

\[
R_{\gamma_n}(f, F_n) = \int_0^\pi \left( \int_0^\pi f(y)dy - F_n(x) \right)^2 dx + \gamma_n \int_0^\pi \left( f^{(p)}(x) \right)^2 dx.
\]

- Given orthonormal functions, \( \psi_1(\theta), \ldots, \psi_k(\theta), \ldots \)

\[
f_n(\theta) = \frac{1}{\pi} + \frac{2}{\pi} \sum_{k=1}^{\infty} \tilde{A}_k \psi_k(\theta),
\]

with \( \tilde{A}_0, \tilde{A}_1, \ldots, \tilde{A}_k, \ldots \) expansion coefficients, \( \tilde{A}_k = \langle f_n, \psi_k \rangle \).

- The coefficients \( \tilde{A}_k \) cannot be directly computed from \( f_n \), but

\[
\tilde{A}_k = \langle f_n, \psi_k \rangle = \langle \hat{f}_n, \hat{\psi}_k \rangle
\]

\[
= \int_0^\pi \left( \frac{1}{1 + \gamma_n u^2 \hat{K}(u)\hat{K}(-u)} \right) \frac{1}{n} \sum_{j=1}^{n} \exp(-iu\theta_j) \cdot \hat{\psi}_k(u)du.
\]
Using cosine series expansions, i.e., \( \psi_k(\theta) = \cos(k\theta) \), it is well-known that

\[
\hat{\psi}_k(u) = \frac{1}{2}(\delta(u - k) + \delta(u + k)).
\]

This facilitates the computation of \( \tilde{A}_k \) avoiding the calculation of the integral. Thus, the minimum of \( R_{\gamma n} \)

\[
\tilde{A}_k = \frac{1}{2n} \left( \frac{1}{1 + \gamma_n(-k)^2 \hat{K}(-k) \hat{K}(k)} \right) \sum_{j=1}^{n} \exp(ik\theta_j)
\]

\[
+ \left( \frac{1}{1 + \gamma_n k^2 \hat{K}(k) \hat{K}(-k)} \right) \sum_{j=1}^{n} \exp(-ik\theta_j)
\]

\[
= \frac{1}{1 + \gamma_n k^2 \hat{K}(k) \hat{K}(-k)} \frac{1}{n} \sum_{j=1}^{n} \cos(k\theta_j) = \frac{1}{1 + \gamma_n k^{2(p+1)}} \frac{1}{n} \sum_{j=1}^{n} \cos(k\theta_j),
\]

where \( \theta_j \in (0, \pi) \) are given samples of the unknown distribution. In the last step, \( \hat{K}(u) = (iu)^p \) is used.
The data-driven COS method

- Employ the solution of the regularization problem for density estimation in the COS framework.
- In both, the density 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, Greeks, etc.
- The samples must follow risk-neutral measure (Monte Carlo paths).
The data-driven COS method

- Key idea: $\tilde{A}_k$ approximates $A_k$.
- Risk neutral samples from an asset at time $T$, $S_1(t), S_2(t), \ldots, S_n(t)$.
- With a logarithmic transformation, we have

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

- The regularization solution is defined in $(0, \pi)$, by transformation

$$\theta_j = \pi \frac{Y_j - a}{b - a},$$

- The boundaries $a$ 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 data-driven COS method - European options

- The $A_k$ coefficients are replaced by the data-driven $\tilde{A}_k$

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

- The ddCOS pricing formula for European options

$$\tilde{\nu}(x, t) = e^{-r(T-t)} \sum_{k=0}^{\infty} \frac{1}{n} \sum_{j=1}^{n} \cos \left( k\pi \frac{Y_j - a}{b-a} \right) \frac{1}{1 + \gamma_n k^2(p+1)} \cdot V_k $$

$$= e^{-r(T-t)} \sum_{k=0}^{\infty} \tilde{A}_k V_k. $$

- As in the original COS method, we must truncate the infinite sum to a finite number of terms $N$

$$\tilde{\nu}(x, t) = e^{-r(T-t)} \sum_{k=0}^{N} \tilde{A}_k V_k, $$
Data-driven expressions for the $\Delta$ and $\Gamma$ sensitivities. Define the corresponding sine coefficients as

$$\tilde{B}_k := \frac{1}{n} \sum_{j=1}^{n} \sin \left( k\pi \frac{Y_j-a}{b-a} \right) \frac{1}{1 + \gamma_n k^2(p+1)}.$$

Taking derivatives of the ddCOS pricing formulat w.r.t the samples, $Y_j$, the data-driven Greeks, $\tilde{\Delta}$ and $\tilde{\Gamma}$, can be obtained by

$$\tilde{\Delta} = e^{-r(T-t)} \sum_{k=0}^{N} \tilde{B}_k \left( - \frac{k\pi}{b-a} \right) \frac{V_k}{S(0)},$$

$$\tilde{\Gamma} = e^{-r(T-t)} \sum_{k=0}^{N} \left( \tilde{B}_k \cdot \frac{k\pi}{b-a} - \tilde{A}_k \left( \frac{k\pi}{b-a} \right)^2 \right) \frac{V_k}{S^2(0)}.$$
Here, we show how to apply *antithetic variates (AV)* to our method. Since the samples must be i.i.d., an immediate application of AV is not possible.

Assume antithetic samples, $Y_i'$, that can be computed without extra computational effort, a new estimator is defined as

$$\tilde{A}_k := \frac{1}{2} \left( \tilde{A}_k + \tilde{A}_k' \right),$$

where $\tilde{A}_k'$ are “antithetic coefficients”, obtained from $Y_i'$. It can be proved that the use of $\tilde{A}_k$ results in a variance reduction.

Additional information to reduce the variance. For example, the martingale property

$$S(T) = S(T) - \frac{1}{n} \sum_{j=1}^{n} S_j(T) + \mathbb{E}[S(T)],$$

$$= S(T) - \frac{1}{n} \sum_{j=1}^{n} S_j(T) + S(0) \exp(rt).$$
The choice of optimal values of $\gamma_n$ and $p$.

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.

For the regularization parameter $\gamma_n$, a rule that ensures asymptotic convergence

$$\gamma_n = \frac{\log \log n}{n}.$$
Applications of the ddCOS method

- Pricing options (no better than Monte Carlo).
- Sensitivities or Greeks.
- Models without analytic characteristic function. SABR model.
- Risk measures: VaR and Expected shortfall.
- Combinations.
Applications of the ddCOS - Option pricing

Figure: Convergence in prices of the ddCOS method: Antithetic Variates (AV); GBM, $S(0) = 100$, $r = 0.1$, $\sigma = 0.3$ and $T = 2$. 

(a) Call: Strike $K = 100$. 
(b) Put: Strike $K = 100$. 

Álvaro Leitao (CWI & TUDelft)
Applications of the ddCOS - Greeks estimation

Figure: Convergence in Greeks of the ddCOS method: Antithetic Variates (AV); GBM, $S(0) = 100$, $r = 0.1$, $\sigma = 0.3$ and $T = 2$.

(a) $\Delta$ (Call): Strike $K = 100$.

(b) $\Gamma$: Strike $K = 100$. 
Applications of the ddCOS - Greeks estimation

<table>
<thead>
<tr>
<th>$K$ (% of $S(0)$)</th>
<th>80%</th>
<th>90%</th>
<th>100%</th>
<th>110%</th>
<th>120%</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ref.</td>
<td>0.8868</td>
<td>0.8243</td>
<td>0.7529</td>
<td>0.6768</td>
<td>0.6002</td>
</tr>
<tr>
<td>ddCOS</td>
<td>0.8867</td>
<td>0.8240</td>
<td>0.7528</td>
<td>0.6769</td>
<td>0.6002</td>
</tr>
<tr>
<td>RE</td>
<td>$1.1012 \times 10^{-4}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCFD</td>
<td>0.8876</td>
<td>0.8247</td>
<td>0.7534</td>
<td>0.6773</td>
<td>0.6006</td>
</tr>
<tr>
<td>RE</td>
<td>$7.5168 \times 10^{-4}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\Gamma$</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref.</td>
<td>0.0045</td>
<td>0.0061</td>
<td>0.0074</td>
<td>0.0085</td>
<td>0.0091</td>
</tr>
<tr>
<td>ddCOS</td>
<td>0.0045</td>
<td>0.0062</td>
<td>0.0075</td>
<td>0.0084</td>
<td>0.0090</td>
</tr>
<tr>
<td>RE</td>
<td>$8.5423 \times 10^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCFD</td>
<td>0.0045</td>
<td>0.0059</td>
<td>0.0071</td>
<td>0.0079</td>
<td>0.0083</td>
</tr>
<tr>
<td>RE</td>
<td>$4.9554 \times 10^{-2}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table: GBM call option Greeks: $S(0) = 100$, $r = 0.1$, $\sigma = 0.3$ and $T = 2$. 
### Applications of the ddCOS - Greeks estimation

<table>
<thead>
<tr>
<th>K (% of S(0))</th>
<th>80%</th>
<th>90%</th>
<th>100%</th>
<th>110%</th>
<th>120%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ref.</td>
<td>0.8385</td>
<td>0.8114</td>
<td>0.7847</td>
<td>0.7584</td>
<td>0.7328</td>
</tr>
<tr>
<td>ddCOS</td>
<td>0.8383</td>
<td>0.8113</td>
<td>0.7846</td>
<td>0.7585</td>
<td>0.7333</td>
</tr>
<tr>
<td>RE</td>
<td>$2.7155 \times 10^{-4}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ref.</td>
<td>0.0022</td>
<td>0.0024</td>
<td>0.0027</td>
<td>0.0029</td>
<td>0.0030</td>
</tr>
<tr>
<td>ddCOS</td>
<td>0.0022</td>
<td>0.0024</td>
<td>0.0027</td>
<td>0.0029</td>
<td>0.0030</td>
</tr>
<tr>
<td>RE</td>
<td>$8.2711 \times 10^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ref.</td>
<td>0.0022</td>
<td>0.0026</td>
<td>0.0028</td>
<td>0.0031</td>
<td>0.0033</td>
</tr>
<tr>
<td>ddCOS</td>
<td>0.0022</td>
<td>0.0026</td>
<td>0.0028</td>
<td>0.0031</td>
<td>0.0033</td>
</tr>
<tr>
<td>RE</td>
<td>$6.118 \times 10^{-2}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table:** Merton jump-diffusion call option Greeks: $S(0) = 100$, $r = 0.1$, $\sigma = 0.3$, $\mu_j = -0.2$, $\sigma_j = 0.2$ and $\lambda = 8$ and $T = 2$. 
### Applications of the ddCOS - Greeks estimation

<table>
<thead>
<tr>
<th>$K$ (% of $S(0)$)</th>
<th>80%</th>
<th>90%</th>
<th>100%</th>
<th>110%</th>
<th>120%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref.</td>
<td>0.9914</td>
<td>0.9284</td>
<td>0.5371</td>
<td>0.0720</td>
<td>0.0058</td>
</tr>
<tr>
<td>ddCOS</td>
<td>0.9916</td>
<td>0.9282</td>
<td>0.5363</td>
<td>0.0732</td>
<td>0.0058</td>
</tr>
<tr>
<td>RE</td>
<td>5.2775 $\times 10^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCFD</td>
<td>0.9911</td>
<td>0.9279</td>
<td>0.5368</td>
<td>0.0737</td>
<td>0.0058</td>
</tr>
<tr>
<td>RE</td>
<td>5.5039 $\times 10^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table: Call option Greek $\Delta$ under the SABR model: $S(0) = 100$, $r = 0$, $\sigma_0 = 0.3$, $\alpha = 0.4$, $\beta = 0.6$, $\rho = -0.25$ and $T = 2$. 
## Applications of the ddCOS - Greeks estimation

<table>
<thead>
<tr>
<th>K (% of $S(0)$)</th>
<th>80%</th>
<th>90%</th>
<th>100%</th>
<th>110%</th>
<th>120%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ref.</strong></td>
<td>0.8384</td>
<td>0.7728</td>
<td>0.6931</td>
<td>0.6027</td>
<td>0.5086</td>
</tr>
<tr>
<td><strong>ddCOS</strong></td>
<td>0.8364</td>
<td>0.7703</td>
<td>0.6902</td>
<td>0.6006</td>
<td>0.5084</td>
</tr>
<tr>
<td><strong>RE</strong></td>
<td>2.7855 $\times 10^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Hagan</strong></td>
<td>0.8577</td>
<td>0.7955</td>
<td>0.7170</td>
<td>0.6249</td>
<td>0.5265</td>
</tr>
<tr>
<td><strong>RE</strong></td>
<td>3.1751 $\times 10^{-2}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table:** $\Delta$ under SABR model. Setting: Call, $S(0) = 0.04$, $r = 0.0$, $\sigma_0 = 0.4$, $\alpha = 0.8$, $\beta = 1.0$, $\rho = -0.5$ and $T = 2$. 
Applications of the ddCOS - Risk measures

- In the context of the Delta-Gamma approach (COS in [OGO14]).
- The change in a portfolio value is defined:
  \[
  L := -\Delta V = V(S, t) - V(S + \Delta S, t + \Delta t).
  \]
- The formal definition of the VaR reads
  \[
  \mathbb{P}(\Delta V < \text{VaR}(q)) = 1 - F_L(\text{VaR}(q)) = q,
  \]
  with \( q \) a predefined confidence level.
- Given the VaR, the ES measure is computed as
  \[
  \text{ES} := \mathbb{E}[\Delta V | \Delta V > \text{VaR}(q)].
  \]
- Two portfolios with the same composition: one European call and half a European put on the same asset, maturity 60 days and \( K = 101 \).
- Different time horizons: 1 day (Portfolio 1) and 10 days (Portfolio 2). The asset follows a GBM with \( S(0) = 100, r = 0.1 \) and \( \sigma = 0.3 \).
Applications of the ddCOS - Risk measures

(a) Density Portfolio 1.

(b) Density Portfolio 2.

Figure: Recovered densities of $L$: ddCOS vs. COS.
Applications of the ddCOS - Risk measures

(a) Portfolio 1: $q = 99\%$.

(b) Portfolio 2: $q = 90\%$.

Figure: VaR and ES convergence in $n$. 
Applications of the ddCOS - Risk measures

- The oscillations can be removed.
- Two options: smoothing parameter or filters [RVO14].

Figure: Smoothed densities of $L$. 

(a) Density Portfolio 1.  
(b) Density Portfolio 2.
Applications of the ddCOS - Risk measures and SABR

(a) VaR and ES: $q = 99\%$.

(b) $F_L$ and $f_L$.

Figure: Delta-Gamma approach under the SABR model. Setting: $S(0) = 100$, $K = 100$, $r = 0.0$, $\sigma_0 = 0.4$, $\alpha = 0.8$, $\beta = 1.0$, $\rho = -0.5$, $T = 2$, $q = 99\%$ and $\Delta t = 1/365$. 
Applications of the ddCOS - Risk measures and SABR

<table>
<thead>
<tr>
<th>q</th>
<th>10%</th>
<th>30%</th>
<th>50%</th>
<th>70%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>VaR</td>
<td>-1.4742</td>
<td>-0.5917</td>
<td>-0.0022</td>
<td>0.5789</td>
<td>1.3862</td>
</tr>
<tr>
<td>ES</td>
<td>0.1972</td>
<td>0.5345</td>
<td>0.8644</td>
<td>1.2517</td>
<td>1.8744</td>
</tr>
</tbody>
</table>

Table: VaR and ES under SABR model. Setting: $S(0) = 100$, $K = 100$, $r = 0.0$, $\sigma_0 = 0.4$, $\alpha = 0.8$, $\beta = 1.0$, $\rho = -0.5$, $T = 2$, and $\Delta t = 1/365$. 
Conclusions

- The ddCOS method extends the COS method applicability to cases when only data samples of the underlying are available.
- The method exploits a closed-form solution, in terms of Fourier cosine expansions, of a regularization problem.
- It allows to develop a data-driven method which can be employed for option pricing and risk management.
- The ddCOS method particularly results in an efficient method for the $\Delta$ and $\Gamma$ sensitivities computation, based solely on the samples.
- It can be employed within the Delta-Gamma approximation for calculating risk measures.
Fang Fang and Cornelis W. Oosterlee. 
A novel pricing method for European options based on Fourier-cosine series expansions. 

Fang Fang and Cornelis W. Oosterlee. 
Pricing early-exercise and discrete barrier options by Fourier-cosine series expansions. 

Luis Ortiz-Gracia and Cornelis W. Oosterlee. 
Efficient VaR and Expected Shortfall computations for nonlinear portfolios within the delta-gamma approach. 

Maria J. Ruijter, Mark Versteegh, and Cornelis W. Oosterlee. 
On the application of spectral filters in a Fourier option pricing technique. 

Vladimir N. Vapnik. 
*Statistical learning theory*. 
Thank you for your attention
Choice of $\gamma_n$

- $\gamma_n$ impacts the efficiency of the ddCOS method: it is related to the number of samples, $n$, and number of terms, $N$.
- For the regularization parameter $\gamma_n$, a rule that ensures asymptotic convergence
  \[
  \gamma_n = \frac{\log \log n}{n}.
  \]
- In practical situations: not optimal.
- Exploit the relation between the empirical and real (unknown) CDFs.
Choice of $\gamma_n$

- 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):

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

- Assume we have an approximation, $F_{\gamma_n}$ (which depends on $\gamma_n$).
- An *almost* optimal $\gamma_n$ 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 $\omega^2$. 
Influence of $\gamma_n$

- To assess the impact of $\gamma_n$: *Mean integrated Squared Error* (MiSE):

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

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

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

- Two main aspects influenced $\gamma_n$: accuracy in $n$ and stability in $N$.
- The quality of the approximated density can be also affected.
Influence of $\gamma_n$

(a) Convergence in terms of $n$

(b) Convergence in terms of $N$

Figure: Influence of $\gamma_n$. 
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:

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

**Data:** $n$, $\gamma_n$

$N_{\text{min}} = 5$

$N_{\text{max}} = \infty$

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

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

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

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

\[
\epsilon_N = \frac{|\text{MiSE}_N - \text{MiSE}_{\text{prev}}|}{|\text{MiSE}_N|}
\]

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

\[
N_{\text{op}} = N
\]

**else**

 Break

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

**Figure:** Almost optimal $N$. 