Efficient Risk Management in Monte Carlo

Module 4: Risk Management by Adjoint Algorithmic Differentiation II

Luca Capriotti

The 7th Summer School in Mathematical Finance
African Institute for Mathematical Sciences, Muizenberg, Cape Town, South Africa,
February 20-22, 2014

Section 1

Outline

Correlation Greeks and Binning Techniques

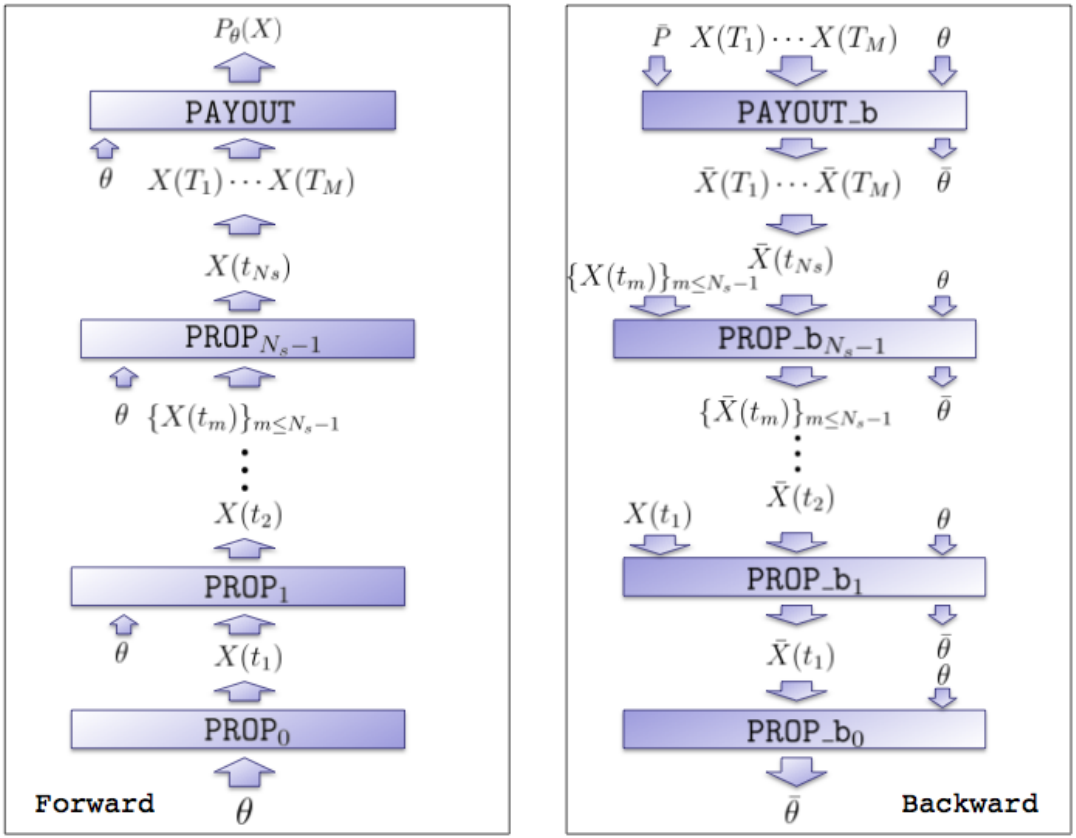
Case Study: Correlation Greeks for Basket Default Contracts

Section 2

Correlation Greeks and Binning Techniques

Correlation Structure of the Random Variates

Recall the general AAD MC design for the computation of the estimators on each MC path:



Correlation Structure of the Random Variates

- In the AAD MC design we have assumed for simplicity that the random variates $Z(t_n)$ entering in the propagation method:

$$X(t_{n+1}) = \texttt{PROP}_n[\{X(t_m)\}_{m \leq n}, Z(t_n), \theta],$$

  are dummy variables carrying no interesting sensitivities.
- As a result, in the corresponding adjoint propagation methods:

$$(\{\bar{X}(t_m)\}_{m \leq n}, \bar{\theta}) \mathrel{+}= \texttt{PROP\_b}_n[\{X(t_m)\}_{m \leq n}, Z(t_n), \theta, \bar{X}(t_{n+1})],$$

  the adjoint of the random variates $\bar{Z}(t_n)$ do not appear among the output.
- If we want to compute the sensitivities with respect to the correlation structure of the random variates, this scheme needs to be extended.

Correlation Structure of the Random Variates

- In a typical setup, the random variates Z_i driving the random processes are correlated.
- For instance, assume that the random variates $Z'(t_n)$ are jointly normal, and denote with $\rho_{ij}(t_m) = \mathbb{E}[Z_i(t_m)Z_j(t_m)]$ the correlation matrix.
- Uncorrelated random variates $Z'(t_n)$ are therefore mapped into their correlated counterparts $Z(t_n)$ and then used to implement the propagation step $X(t_n) \to X(t_{n+1})$ so that the propagation step is modified as

$$\begin{aligned} Z(t_n) &= \texttt{CORRELATE}(Z'(t_n), \theta) \\ X(t_{n+1}) &= \texttt{PROP}_n[\{X(t_m)\}_{m \leq n}, Z(t_n), \theta] \ , \end{aligned}$$

where we have included the correlation parameters defining the correlation matrix ρ in the vector θ.

Modified Adjoint of the Propagation Step

- The adjoint of the Propagation Step

$$X(t_{n+1}) = \texttt{PROP}_n[\{X(t_m)\}_{m\leq n}, Z(t_n), \theta],$$

is modified as

$$(\{\bar{X}(t_m)\}_{m\leq n}, \bar{\theta}, \bar{Z}(t_n)) \mathrel{+}= \texttt{PROP\_b}_n[\{X(t_m)\}_{m\leq n}, Z(t_n), \theta, \bar{X}(t_{n+1})],$$

where

$$\bar{X}(t_m) \mathrel{+}= \sum_{j=1}^{N} \bar{X}_j(t_{n+1}) \frac{\partial X_j(t_{n+1})}{\partial X(t_m)} \qquad \bar{\theta} \mathrel{+}= \sum_{j=1}^{N} \bar{X}_j(t_{n+1}) \frac{\partial X_j(t_{n+1})}{\partial \theta},$$

with $m = 1, \ldots, n$. Here the additional output is given by the adjoint of the correlated variates:

$$\bar{Z}(t_n) \mathrel{+}= \sum_{j=1}^{N} \bar{X}_j(t_{n+1}) \frac{\partial X_j(t_{n+1})}{\partial Z(t_n)}.$$

Adjoint of the Correlation Step

- The adjoint of the Correlation Step

$$Z(t_n) = \texttt{CORRELATE}(Z'(t_n), \theta),$$

reads

$$\bar{\theta} \mathrel{+}= \texttt{CORRELATE\_b}(Z'(t_n), \theta, \bar{Z}(t_n)),$$

corresponding to the operation

$$\bar{\theta} \mathrel{+}= \sum_{j=1}^{N} \bar{Z}'_j(t_n) \frac{\partial Z_j(t_n)}{\partial \theta}$$

updating the components of the vector θ corresponding to the adjoint of the correlation parameters.

Example: Cholesky Factorization

- In a simple setup the method `CORRELATE` generally involves the so-called Cholesky factorization of an $N \times N$ correlation matrix ρ.
- Recall that the Cholesky factorization of a Hermitian positive-definite matrix ρ produces a lower triangular $N \times N$ matrix L such that $\rho = LL^T$.
- Given the Cholesky factor L, and a vector of N uncorrelated normal Z', it is immediate to verify that $Z = LZ'$ are correlated normal such that $\mathbb{E}[Z_i Z_j] = \rho_{ij}$

Adjoint of the Cholesky Factorization

- When implemented in terms of the Cholesky factorization, the method `CORRELATE` reads

  Step 1 Perform Cholesky factorization, say $L = \texttt{CHOLESKY}(\rho)$.
  Step 2 Compute: $Z = LZ'$.

- The corresponding method `CORRELATE_b` reads

  Step $\bar{2}$ Compute: $\bar{L} = \bar{Z} Z'^{t}$.
  Step $\bar{1}$ Compute: $\bar{\rho} = \texttt{CHOLESKY\_b}(\rho, \bar{L})$, where

$$\bar{\rho}_{i,j} = \sum_{l,m=1}^{N} \frac{\partial L_{l,m}}{\partial \rho_{i,j}} \bar{L}_{l,m},$$

  providing the sensitivities with respect to the entries of the correlation matrix. These are copied in the appropriate components of the vector $\bar{\theta}$.

- Note that Z' are now dummy integration variables (sampled stochastically). Thefore their adjoints $\bar{Z}'$ are not computed.

Adjoint of the Cholesky Factorization (Pseudocode)

```
Cholesky_b(rho, L_b,rho_b)

   // Forward Sweep
   for (i=0 .. n-1)
      for (j=i .. n-1)
      sum[i,j] = rho[i,j];
      for (k=i-1 .. 0)
         sum[i,j] -= L[i,k] * L[j,k];
         if (i == j)
           L[i,i] = sqrt(sum[i,j]);
         else
           L[j,i] = sum[i,j] / L[i,i];

   // Backward Sweep
   for (i=n-1 .. 0)
      for (j=n-1 .. i)
         sum_b =  0.0;

         if (i == j)
           if (sum[i,j] == 0.0)
              sum_b = 0.0;
           else
              sum_b = L_b[i,j]/( 2.0 * L[i,j]);
           L_b[i,j] = 0.0;
         else
           sum_b = L_b[j,i]/L[i,i];
           L_b[i,i] -= sum[i,j] * sum_b / L[i,i];
           L_b[j,i] = 0.0;

         for (k=i-1 .. 0)
           L_b[i,k] -= L[j,k]*sum_b;
           L_b[j,k] -= L[i,k]*sum_b;

         rho_b[i,j] += sum_b;
```

The adjoint algorithm contains the original Cholesky factorization plus a backward sweep with the same complexity and a similar number of operations.

Hence, as expected, the computational cost is just a small multiple (of order 2, in this case) of the cost of evaluating the original factorization.

Adjoint of the Cholesky Factorization

- The Cholesky factorization $L = \texttt{CHOLESKY}(\rho)$ does not depend on the random variates Z therefore it can be performed before the first Monte Carlo path is performed. As a result, `CORRELATE` consists of the matrix multiplication $Z = LZ'$, only.

- Similarly the Adjoint of `CORRELATE_b` consists only of the step $\bar{L} = \bar{Z}'Z^t$, ($\bar{\theta}$ will contain the adjoint of the Cholesky factors L rather than the entries of the correlation matrix ρ) and the Adjoint of the Cholesky factorization

$$\bar{\rho} = \texttt{CHOLESKY\_b}(\rho, \bar{L})$$

can be performed after the end of the backward sweep after the last MC path..

Statistical Uncertainties

- Given the MC estimators for the Cholesky factors sensitivities $\langle \bar{L} \rangle = \langle \partial V(X)/\partial L \rangle$ and their statistical uncertainties

$$\langle \bar{L} \rangle = \frac{1}{N_{\mathrm{MC}}} \sum_{i_{\mathrm{MC}}=1}^{N_{\mathrm{MC}}} \bar{L}(X[i_{\mathrm{MC}}]) \qquad \sigma_{\bar{L}} = \sqrt{\frac{1}{N_{\mathrm{MC}}} \sum_{i_{\mathrm{MC}}=1}^{N_{\mathrm{MC}}} \left(\bar{L}(X[i_{\mathrm{MC}}])^2 - \langle \bar{L} \rangle \right)^2}$$

- One can compute the estimator for the correlation sensitivities via the Cholsesky factorization

$$\langle \bar{\rho} \rangle = \texttt{CHOLESKY\_b}(\rho, \langle \bar{L} \rangle)$$

but not their sensitivities:

$$\sigma_{\bar{\rho}} \neq \texttt{CHOLESKY\_b}(\rho, \sigma_{\bar{L}})$$

- Performing the adjoint of the Cholesky decomposition once per simulation does not allow the calculation of a confidence interval for the correlation sensitivities.

Path by Path Adjoint Cholesky Factorization

- An alternative approach would be to convert $\bar{L}$ to $\bar{\rho}$ for each individual path $i_{\mathrm{MC}} = 1, \ldots, N_{\mathrm{MC}}$

$$\bar{\rho}(X[i_{\mathrm{MC}}]) = \texttt{CHOLESKY\_b}(\rho, \bar{L}(X[i_{\mathrm{MC}}]))$$

and then compute the average and standard deviation of $\bar{\rho}[i_{\mathrm{MC}}]$ in the usual way:

$$\langle \bar{\rho} \rangle = \frac{1}{N_{\mathrm{MC}}} \sum_{i_{\mathrm{MC}}=1}^{N_{\mathrm{MC}}} \bar{\rho}(X[i_{\mathrm{MC}}]) \qquad \sigma_{\bar{\rho}} = \sqrt{\frac{1}{N_{\mathrm{MC}}} \sum_{i_{\mathrm{MC}}=1}^{N_{\mathrm{MC}}} \left(\bar{\rho}(X[i_{\mathrm{MC}}])^2 - \langle \bar{\rho} \rangle \right)^2}$$

However, this is rather costly.

Binning

- An excellent compromise between these two extremes is to divide the N_{MC} paths into N_{B} 'bins' of equal size $n = N/N_{\mathrm{B}}$.
- For each bin $j_{\mathrm{B}} = 1, \ldots, N_{\mathrm{B}}$, an average value of $\langle \bar{L} \rangle_{j_{\mathrm{B}}}$ is computed

$$\langle \bar{L} \rangle_{j_{\mathrm{B}}} = \frac{1}{n} \sum_{i_{\mathrm{MC}}=1}^{n} \bar{L}(X[i_{\mathrm{MC}}])$$

and converted into a corresponding value for

$$\langle \bar{\rho} \rangle_{j_{\mathrm{B}}} = \texttt{CHOLESKY\_b}(\rho, \langle \bar{L} \rangle_{j_{\mathrm{B}}}).$$

Binning

- These N_b estimates for $\bar{\rho}$ can then be combined in the usual way to form an overall estimate of the correlation risk:

$$\langle \bar{\rho} \rangle = \frac{1}{N_{\mathrm{B}}} \sum_{j_{\mathrm{B}}=1}^{N_{\mathrm{B}}} \langle \bar{\rho} \rangle_{j_{\mathrm{B}}} = \frac{1}{N_{\mathrm{MC}}} \sum_{i_{\mathrm{MC}}=1}^{N_{\mathrm{MC}}} \bar{\rho}(X[i_{\mathrm{MC}}]),$$

  where the second equality follows from the linearity of the adjoint functions, and the associated confidence interval:

$$\sigma_{\bar{\rho}} = \sqrt{\frac{1}{N_{\mathrm{B}}} \sum_{j_{\mathrm{B}}=1}^{N_{\mathrm{B}}} \left(\langle \bar{\rho} \rangle_{j_{\mathrm{B}}}^2 - \langle \bar{\rho} \rangle \right)^2}.$$

Binning

- In the standard evaluation, the cost of the Cholesky factorization is $O(N^3)$, and the cost of the MC sampling is $O(N_{\mathrm{MC}}N^2)$, so the total cost is $O(N^3 + N_{\mathrm{MC}}N^2)$. Since N_{MC} is always much greater than N, the cost of the Cholesky factorization is usually negligible.
- The cost of the adjoint steps in the MC sampling is also $O(N_{\mathrm{MC}}N^2)$, and when using N_b bins the cost of the adjoint Cholesky factorization is $O(N_{\mathrm{B}}N^3)$.
- To obtain an accurate confidence interval, but with the cost of the Cholesky factorisation being negligible, requires that N_{B} is chosen so that $1 \ll N_{\mathrm{B}} \ll N_{\mathrm{MC}}/N$.
- Without binning, i.e., using $N_{\mathrm{B}} = N_{\mathrm{MC}}$, the cost to calculate the average of the estimators for $\langle\rho\rangle$ is $O(N_{\mathrm{MC}}N^3)$, and so the relative cost compared to the evaluation of the option value is $O(N)$.

Binning and Risk Transforms

- We have presented Binning in the context of the calculation of correlation risk, but there is nothing specific to correlation. In fact these ideas can be applied everytime some computational preprocessing is performed before the MC simulation, and we need to transform the adjoint MC estimators and their confidence interval into the corresponding quantities for the inputs of such preprocessing.
- This is the case for instance when a calibration routine performed before the MC simulation transforms some market inputs $M = (M_1, \ldots, M_{N_M})$, corresponding to the observable prices of securities which the model is calibrated to, into the set of internal model parameters that are used in the MC simulation θ:

$$\theta = \texttt{CALIBRATION}(M).$$

Binning and Risk Transforms

- The binned MC estimators of the adjoint of the internal model parameters $\langle \bar{\theta} \rangle_{j_\mathrm{B}}$ can be transformed into binned MC estimators of the market inputs

$$\langle \bar{M} \rangle_{j_\mathrm{B}} = \texttt{CALIBRATION\_B}(M, \langle \bar{\theta} \rangle_{j_\mathrm{B}}).$$

- Then their distribution can be used to construct the overall MC estimator and the associated statistical uncertainty

$$\langle \bar{M} \rangle = \frac{1}{N_\mathrm{B}} \sum_{j_\mathrm{B}=1}^{N_\mathrm{B}} \langle \bar{M} \rangle_{j_\mathrm{B}},$$

$$\sigma_{\bar{M}} = \sqrt{\frac{1}{N_\mathrm{B}} \sum_{j_\mathrm{B}=1}^{N_\mathrm{B}} \left(\langle \bar{M} \rangle_{j_\mathrm{B}}^2 - \langle \bar{M} \rangle \right)^2}.$$

Section 3

Case Study: Correlation Greeks for Basket Default Contracts

Credit Basket Contracts

- Credit basket contracts are derivatives that are contingent on credit events (defaults for short) of a pool of reference entities typically sovereign, financial or corporate. Generally the credit event is defined as failure to pay a specific liability, say a coupon on a specific bond or category of bonds referenced by the contract, but it can include other events not involving a proper default, like a restructuring of the debt, or regulatory action on a financial institution.
- n-th to default, Collateralized Debt Obligations (CDO) and their variations are examples of credit basket products.
- In the context of basket credit default products the random factors X_i are the time of default τ_i of the i-the reference entity in a basket of N names and the payoff is of the form:

$$P = P(\tau_1, \ldots, \tau_N)$$

Example: n-th to default Basket Default Swap

- In a n-th to default Basket Default Swap one party (protection buyer) makes regular payments to a counterparty (protection seller) at time $T_1, \ldots, T_M \leq T$ provided that less than n defaults events among the components of the basket are observed before time T_M.
- If n defaults occur before time T, the regular payments cease and the protection seller makes a payment to the buyer of $(1 - R_i)$ per unit notional, where R_i is the normalized recovery rate of the i-th asset.
- The value at time zero of the Basket Default Swap on a given realization of the default times $\tau_1, \ldots, \tau_N$, i.e., the Payout function, can be expressed as

$$P(\tau_1, \ldots, \tau_N) = P_{prot}(\tau_1, \ldots, \tau_N) - P_{prem}(\tau_1, \ldots, \tau_N)$$

i.e., as the difference between the so-called *protection* and *premium* legs.

Example: n-th to default Basket Default Swap

- The value leg is given by

$$P_{prot}(\tau_1, \ldots, \tau_N) = (1 - R_n) D(\tau) \mathbb{I}(\tau \leq T),$$

where R_n and τ are the recovery rate and default time of the n-th to default, respectively, $D(t)$ is the discount factor for the interval $[0, t]$ (here we assume for simplicity uncorrelated default times and interest rates), and $\mathbb{I}(\tau \leq T)$ is the indicator function of the event that the n-th default occurs before T.

- The premium leg reads instead, neglecting for simplicity any accrued payment,

$$P_{prem}(\tau_1, \ldots, \tau_N) = \sum_{k=1}^{T_M} s_k D(T_k) \mathbb{I}(\tau \geq T_k)$$

where c_k is the premium payment (per unit notional) at time T_k.

Copula Models

- Credit Basket Products are also known as *correlation products* because their value depends not only on the marginal distribution of the default times but also on their correlation structure.
- Such correlation structure is typically captured by means of a copula model. For instance, in a Gaussian copula, the cumulative joint distribution of default times is assumed of the form:

$$\mathbb{P}(\tau_1 \leq t_1, \ldots, \tau_N \leq t_N) = \Phi_N(\Phi^{-1}(F_1(t_1)), \ldots, \Phi^{-1}(F_N(t_N)); \rho)$$

where $\Phi_N(Z_1, \ldots, Z_N; \rho)$ is a N-dimensional multivariate Gaussian distribution with zero mean, and a $N \times N$ positive semidefinite correlation matrix ρ; Φ^{-1} is the inverse of the standard normal cumulative distribution, and $F_i(t) = \mathbb{P}(\tau_i \leq t)$, $i = 1, \ldots, N$, are the marginal distributions of the default times of each reference entity, depending on a set of model parameters θ.

Hazard Rate Model

- The key concept for the valuation of credit derivatives, in the context of the models generally used in practice, is the *hazard rate*, λ_u, representing the probability of default of the reference entity between times u and $u + du$, conditional on survival up to time u. The hazard rate function λ_u i s commonly parameterized as piece-wise constant with M knot points at time $(t_1, \ldots, t_M)$, $\lambda = (\lambda_1, \ldots, \lambda_M)$, .
- By modelling the default event of a reference entity i as the first arrival time of a Poisson process with intensity λ_u^i, the survival probability, $\mathbb{P}(\tau_i > t)$, is given by

$$\mathbb{P}(\tau_i > t) = \exp\left[-\int_0^t du\ \lambda_u\right],$$

so that the marginal cumulative distribution of default times reads

$$F_i(t; \lambda^i) = \mathbb{P}(\tau \leq t) = 1 - \exp\left[-\int_0^t du\ \lambda_u^i\right],$$

Forward Simulation Algorithm

The simulation of a Gaussian Copula model can be seen as a single time-step instance of the general approach, consisting of the following steps:

Step 0 Perform a Cholesky factorization of the matrix ρ, say $L = \texttt{CHOLESKY}(\rho)$.

For each MC replication:

Step 1 Generate a N dimensional vector of uncorrelated normal Gaussian variates Z'.

Step 2 Correlate the random variates: $Z = \texttt{CORRELATE}(Z', L)$, where as previously discussed the correlation step consist of a single matrix vector multiplication $Z = LZ'$.

Step 3 Perform the 'propagation step' $\tau = \texttt{PROP}_0[Z, \theta]$.

Step 4 Evaluate the payout function: $P = P(\tau)$.

Forward Simulation Algorithm

- From the form of the cumulative joint distribution of default times

$$\mathbb{P}(\tau_1 \leq t_1, \ldots, \tau_N \leq t_N) = \Phi_N(\Phi^{-1}(F_1(t_1; \lambda^1)), \ldots, \Phi^{-1}(F_N(t_N, \lambda^N)); \rho)$$

  it follows that the random variates $\Phi^{-1}(F_1(\tau_i, \lambda^i))$ are distributed according to a multivariate normal distribution.
- Hence the propagation step $\tau = \texttt{PROP}_0[Z, \theta]$ consists in turn of the following sub-steps:

  Step 3a Set $U_i = \Phi(Z_i)$, $i = 1, \ldots, N$.
  Step 3b Set $\tau_i = F_i^{-1}(U_i; \lambda_i)$, $i = 1, \ldots, N$.

  where $F_i^{-1}(U_i; \lambda_i)$ is the root τ_i of the equation

$$\exp\left[-\int_0^{\tau_i} du\ \lambda_u^i\right] = 1 - U_i.$$

Adjoint Simulation Algorithm

- The corresponding adjoint algorithm consists of the following steps:

  Step $\bar{4}$ Evaluate the adjoint Payout $\bar{\tau}_i = \partial P / \partial \tau_i$, for $i = 1, \ldots, N$.
  Step $\bar{3}$ Evaluate the adjoint of the propagation step:

$$(\bar{\lambda}, \bar{Z}) = \texttt{PROP\_b}_0[Z, \theta, \bar{\tau}]$$

  Step $\bar{2}$ Calculate the adjoint of the correlation step:

$$\bar{L} = \texttt{CORRELATE\_b}(Z', \bar{Z}),$$

  implemented as

$$\bar{L} = \bar{Z} Z'^t.$$

Adjoint Simulation Algorithm

- In turn, the adjoint of the corelation step reads:

  Step $\bar{3}b$ Calculate:

$$\bar{U}_i = \bar{\tau}_i \frac{\partial F_i^{-1}(U_i;\lambda^i)}{\partial U^i} = \bar{\tau}_i \frac{1}{f_i(F_i^{-1}(U_i;\lambda^i);\lambda)},$$

$$\bar{\lambda}_j^i = \bar{\tau}_i \frac{\partial F_i^{-1}(U_i;\lambda^i)}{\partial \lambda_j^i},$$

  for $i = 1,\ldots,N$ and $j = 1,\ldots,M$.

  Step $\bar{3}a$ Calculate: $\bar{Z}_i = \bar{U}_i \phi(Z_i)$, $i = 1,\ldots,N$.

where $f_i(t;\lambda) = \partial F(t;\lambda)/\partial t$ is the p.d.f. of the default time of the i-th reference entity and $\phi(x)$ is the standard normal p.d.f. Note that computing the derivative $\partial F_i^{-1}(U_i;\lambda^i)/\partial \lambda_j^i$ involves differentiating the root searching algorithm used to determine the default time τ_i. However, a much better implementation is possible by means of the so-called implicit function theorem [1].

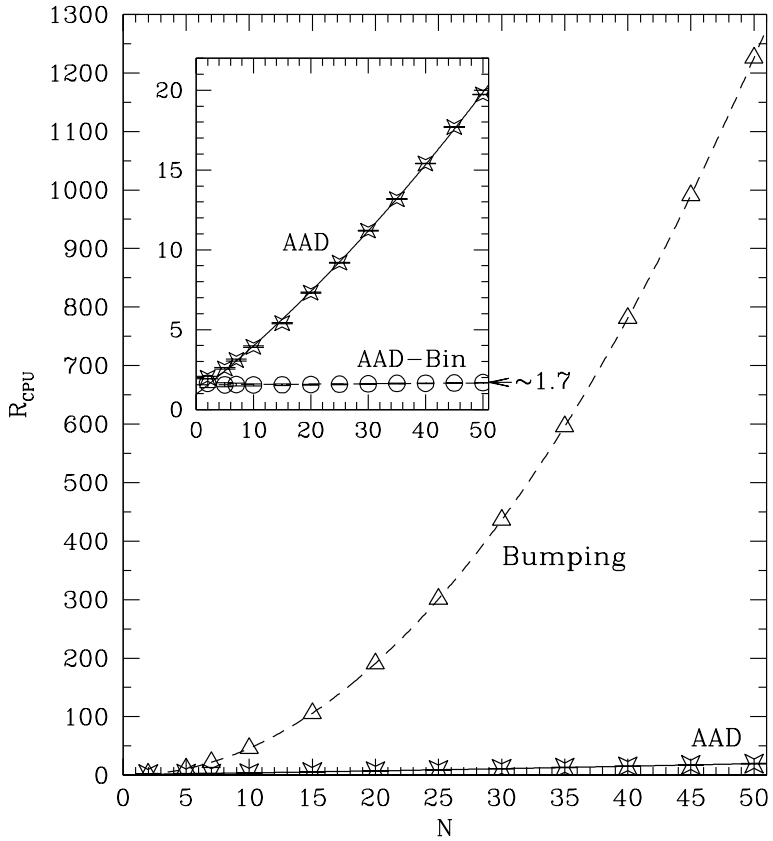
Payout Smoothiong

- In order to apply the Pathwise Derivative method to the payout above, the indicator functions in the premium and protection legs

$$P_{prem}(\tau_1, \ldots, \tau_N) = \sum_{k=1}^{T_M} s_k D(T_k) \mathbb{I}(\tau \geq T_k),$$
$$P_{prot}(\tau_1, \ldots, \tau_N) = (1 - R_n) D(\tau) \mathbb{I}(\tau \leq T),$$

  need to be regularized.
- As seen before, one simple and practical way of doing that is to replace the indicator functions with their smoothed counterpart, at the price of introducing a small amount of bias in the Greek estimators.
- For the problem at hand, as it is also generally the case, such bias can be easily reduced to be smaller than the statistical errors that can be obtained for any realistic number of MC iteration N_{MC}.

Results



Ratios of the CPU time required for the calculation of the option value, and correlation Greeks, and the CPU time spent for the computation of the value alone, as functions of the number of names in the basket. Symbols: Bumping (one-sided finite differences) (triangles), AAD without binning (i.e. $N_{\mathrm{B}} = N_{\mathrm{MC}}$) (stars), AAD with binning ($N_{\mathrm{B}} = 20$) (empty circles).

Results

- As expected, for standard finite-difference estimators, such ratio increases quadratically with the number of names in the basket. Already for medium sized basket ($N \simeq 20$) the cost associated with Bumping is over 100 times more expensive than the one of AAD.
- Nevertheless, at a closer look (see the inset) the relative cost of AAD without binning is $O(N)$, because of the contribution of the adjoint of the Cholsesky decomposition.
- However, when using $N_{\mathrm{B}} = 20$ bins the cost of the adjoint Cholesky computation is negligible and the numerical results show that all the Correlation Greeks can be obtained with a mere 70% overhead compared to the calculation of the value of the option.
- This results in over 2 orders of magnitude savings in computational time for a basket of over 40 Names.

References I

[1] L. Capriotti and J. Lee, *Adjoint Credit Risk Management*, preprint available on SSRN, (2013).

[2] S. P. Smith, *Differentiation of Cholesky Algorithm*, J. of Computational and Graphic Statistics, **4**, 134 (1995).

[3] L. Capriotti and M. Giles, *Fast Correlation Greeks by Adjoint Algorithmic Differentiation*, Risk, April (2010).

[4] L. Capriotti and M. Giles, *Algorithmic Differentiation: Adjoint Greeks Made Easy*, Risk, September (2012).

See also:

- My Publications' Page