Rolling Adjoints

Fast Greeks along Monte Carlo scenarios for early-exercise options

Shashi Jain, Álvaro Leitao and Cornelis W. Oosterlee

QuantMinds International - Lisbon

May 16, 2018

Motivation

- Efficient calculation of option sensitivities is a problem of practical importance.
- For many pricing problems, Monte Carlo is the only feasible choice, as typically for early-exercise options.
- Usual finite differences approach (*bump-and-revalue*) provides poor estimations at high computational cost.
- Sensitivities along the paths, i.e. at intermediate times, is even more involved.
- "Generalization" of the *Smoking adjoints* technique by Giles and Glasserman to a generic interval.
- Sensitivities required for MVA calculations.
- Hedging in energy markets: multiple exercise contracts.

Outline

Problem formulation

- $d-$dimensional Bermudan option pricing problem.
- $\mathbf{X}_t = (X_t^1, \ldots, X_t^d) \in \mathbb{R}^d$, depending on parameters $\theta = \{\theta_1, \ldots, \theta_{N_\theta}\}$.
- Let $h_t := h(\mathbf{X}_t)$ the intrinsic value of the option at time t.
- The holder receives $\max(h_t, 0)$, if the option is exercised.
- The problem is to compute

$$\frac{V_{t_0}(\mathbf{X}_{t_0})}{B_{t_0}} = \max_{\tau} \mathbb{E}\left[\frac{h(\mathbf{X}_\tau)}{B_\tau}\right],$$

  where B_t is the risk-free saving account process and τ is a stopping time.
- Optimization problem: determine the early-exercise policy.

Problem formulation

- It can be solved by the dynamic programming principle.
- The option value at the terminal time T is

$$V_T(\mathbf{X}_T) = \max(h(\mathbf{X}_T), 0).$$

- We solve the problem recursively, moving backwards in time.
- The continuation value $Q_{t_{m-1}}$ is given by

$$Q_{t_{m-1}}(\mathbf{X}_{t_{m-1}}) = B_{t_{m-1}} \mathbb{E}\left[\frac{V_{t_m}(\mathbf{X}_{t_m})}{B_{t_m}} \middle| \mathbf{X}_{t_{m-1}} \right].$$

- The Bermudan option value at time t_{m-1} and state $\mathbf{X}_{t_{m-1}}$ reads

$$V_{t_{m-1}}(\mathbf{X}_{t_{m-1}}) = \max(h(\mathbf{X}_{t_{m-1}}), Q_{t_{m-1}}(\mathbf{X}_{t_{m-1}})).$$

- We are interested in V_{t_0}.

Stochastic Grid Bundling Method (SGBM)

- SGBM is based on N independent paths, $\{\mathbf{X}_{t_0}, \dots, \mathbf{X}_{t_M}\}$, obtained by a discretization scheme

$$\mathbf{X}_{t_m}(n) = F_{m-1}(\mathbf{X}_{t_{m-1}}(n), \mathbf{Z}_{t_{m-1}}(n), \theta),$$

  where $n = 1, \dots, N$ is the index of the path.
- $\mathbf{Z}_{t_{m-1}}$ is a d-dimensional standard normal random vector.
- F_{m-1} is a transformation from $\mathbb{R}^d$ to $\mathbb{R}^d$.
- The method starts by computing the option value at terminal time as

$$V_{t_M}(\mathbf{X}_{t_M}) = \max(h(\mathbf{X}_{t_M}), 0).$$

- The following SGBM components are performed for each time step, t_m, $m \leq M$, moving backwards in time, starting from t_M.

SGBM - Bundling

- The grid points at t_{m-1} are *bundled* into $\mathcal{B}_{t_{m-1}}(1), \ldots, \mathcal{B}_{t_{m-1}}(\nu)$ non-overlapping sets or partitions.
- Several bundling techniques can be employed,
  - *Equal-partitioning*
  - *k-means clustering algorithm*
  - *recursive bifurcation*
  - *recursive bifurcation of a reduced state space*
- A mapping $\mathcal{I}_{t_{m-1}}^{\beta} : \mathbb{N}^{[1,N_\beta]} \mapsto \mathbb{N}^{[1,N]}$, is defined which maps ordered indices of paths in a bundle $\mathcal{B}_{t_{m-1}}(\beta)$ to the original path indices, where $N_\beta := |\mathcal{B}_{t_{m-1}}(\beta)|$ is the cardinality of the β-th bundle, $\beta = 1, \ldots, \nu$.

SGBM - Regression

- **Regress-later** approach within each bundle $\mathcal{B}_{t_{m-1}}(\beta),\ \beta = 1, \dots, \nu$.
- A parameterized value function $\tilde{G} : \mathbb{R}^d \times \mathbb{R}^K \mapsto \mathbb{R}$, which assigns values $\tilde{G}(\mathbf{X}_{t_m}, \alpha^{\beta}_{t_m})$ to states $\mathbf{X}_{t_m}$, is introduced.
- The aim is to choose, for each t_m and β, a vector $\alpha^{\beta}_{t_m}$ so that

$$\tilde{G}\left(\mathbf{X}_{t_m}, \alpha^{\beta}_{t_m}\right) = V_{t_m}(\mathbf{X}_{t_m}).$$

- The option value is approximated as a linear combination of a finite number of orthonormal basis functions ϕ_k as

$$V_{t_m}(\mathbf{X}_{t_m}) \approx \widehat{G}\left(\mathbf{X}_{t_m}, \alpha^{\beta}_{t_m}\right) := \sum_{k=1}^{K} \alpha^{\beta}_{t_m}(k)\phi_k(\mathbf{X}_{t_m}).$$

- The $\alpha^{\beta}_{t_m}$ weights are approximated using a least squares regression by

$$\underset{\widehat{\alpha}^{\beta}_{t_m}}{\operatorname{argmin}} \sum_{n=1}^{N_\beta} \left(V_{t_m}\left(\mathbf{X}_{t_m}\left(\mathcal{I}^{\beta}_{t_{m-1}}(n)\right)\right) - \sum_{k=1}^{K} \widehat{\alpha}^{\beta}_{t_m}(k)\phi_k\left(\mathbf{X}_{t_m}\left(\mathcal{I}^{\beta}_{t_{m-1}}(n)\right)\right)\right)^2.$$

SGBM - Continuation and option values

- The continuation values for $\mathbf{X}_{t_{m-1}}(n) \in \mathcal{B}_{t_{m-1}}(\beta)$, $n = 1, \ldots, N$, $\beta = 1, \ldots, \nu$, are approximated by

$$\widehat{Q}_{t_{m-1}}\left(\mathbf{X}_{t_{m-1}}(n)\right) = \mathbb{E}\left[\widehat{G}\left(\mathbf{X}_{t_m}, \alpha_{t_m}^{\beta}\right) \mid \mathbf{X}_{t_{m-1}}(n)\right].$$

- Exploiting the linearity of the expectation operator, it is written as

$$\widehat{Q}_{t_{m-1}}(\mathbf{X}_{t_{m-1}}(n)) = \sum_{k=1}^{K} \widehat{\alpha}_{t_m}^{\beta}(k)\mathbb{E}\left[\phi_k(\mathbf{X}_{t_m}) \mid \mathbf{X}_{t_{m-1}}(n)\right].$$

- The vector of basis functions ϕ_k should ideally be chosen such that the expectations $\mathbb{E}\left[\phi_k(\mathbf{X}_{t_m})|\mathbf{X}_{t_{m-1}}\right]$ are known in closed-form, or have analytic approximations.

- The option value at each exercise time is then given by

$$\widehat{V}_{t_{m-1}}\left(\mathbf{X}_{t_{m-1}}(n)\right) = \max\left(h\left(\mathbf{X}_{t_{m-1}}(n)\right), \widehat{Q}_{t_{m-1}}\left(\mathbf{X}_{t_{m-1}}(n)\right)\right).$$

Sensitivities along the paths with SGBM

- Naturally, we follow a backward iteration, starting at maturity, where the sensitivities are again trivial to calculate.
- We focus on two main sensitivities of interest:
  - With respect to $\mathbf{X}_{t_{m-1}}$, i.e. $\frac{\partial V_{t_{m-1}}(\mathbf{X}_{t_{m-1}})}{\partial \mathbf{X}_{\mathbf{t_{m-1}}}}$.
  - With respect to the model parameters, $\frac{\partial V_{t_{m-1}}(\mathbf{X}_{t_{m-1}})}{\partial \theta}$.
- The method requires the derivatives of the regression coefficients, $\widehat{\alpha}_{t_m}^{\beta}$.
- Assuming minimal smoothness of the option value function V,

$$\frac{\partial}{\partial \theta}\left(\mathbb{E}\left[\frac{V_{t_m}(\mathbf{X}_{t_m})}{B_{t_m}}\bigg|\mathbf{X}_{t_{m-1}}\right]\right) = \mathbb{E}\left[\frac{\partial}{\partial \theta}\left(\frac{V_{t_m}(\mathbf{X}_{t_m})}{B_{t_m}}\right)\bigg|\mathbf{X}_{t_{m-1}}\right].$$

Delta along the paths

- Delta is the sensitivity of the option value at t_{m-1} w.r.t. $\mathbf{X}_{t_{m-1}}$,

$$\begin{aligned}\frac{\partial V_{t_{m-1}}(\mathbf{X}_{t_{m-1}})}{\partial \mathbf{X}_{t_{m-1}}} &= \left(\frac{\partial h\left(\mathbf{X}_{t_{m-1}}\right)}{\partial \mathbf{X}_{t_{m-1}}}\right)\mathbb{1}_{Q_{t_{m-1}}<h(\mathbf{X}_{t_{m-1}})} \\ &+ \left(\frac{\partial Q_{t_{m-1}}\left(\mathbf{X}_{t_{m-1}}\right)}{\partial \mathbf{X}_{t_{m-1}}}\right)\mathbb{1}_{Q_{t_{m-1}}\geq h(\mathbf{X}_{t_{m-1}})}.\end{aligned}$$

- The derivative of the immediate payoff, h, is usually easy to compute.
- The computation of the sensitivity of the continuation value function

$$\begin{aligned}\frac{\partial \widehat{Q}_{t_{m-1}}(\mathbf{X}_{t_{m-1}}(n))}{\partial \mathbf{X}_{t_{m-1}}} &= \frac{\partial}{\partial \mathbf{X}_{t_{m-1}}}\left(\sum_{k=1}^{K}\widehat{\alpha}^{\beta}_{t_m}(k)\mathbb{E}\left[\phi_k(\mathbf{X}_{t_m}) \mid \mathbf{X}_{t_{m-1}}(n)\right]\right) \\ &= \sum_{k=1}^{K}\left(\frac{\partial \widehat{\alpha}^{\beta}_{t_m}(k)}{\partial \mathbf{X}_{t_{m-1}}}\mathbb{E}\left[\phi_k(\mathbf{X}_{t_m}) \mid \mathbf{X}_{t_{m-1}}(n)\right]\right. \\ &+ \left.\widehat{\alpha}^{\beta}_{t_m}(k)\frac{\partial}{\partial \mathbf{X}_{t_{m-1}}}\mathbb{E}\left[\phi_k(\mathbf{X}_{t_m}) \mid \mathbf{X}_{t_{m-1}}(n)\right]\right).\end{aligned}$$

Delta along the paths

- $\frac{\partial}{\partial \mathbf{X}_{t_{m-1}}} \mathbb{E}\left[\phi_k(\mathbf{X}_{t_m}) \mid \mathbf{X}_{t_{m-1}}(n)\right]$ is readily computed.
- The derivative of the regression coefficients is the difficult part.
- Let us first define matrix $\mathbf{A}^{\beta}_{t_m}$ as

$$\mathbf{A}^{\beta}_{t_m} := \begin{bmatrix} \phi_1(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(1))) & \phi_2(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(1))) & \dots & \phi_K(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(1))) \\ \phi_1(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(2))) & \phi_2(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(2))) & \dots & \phi_K(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(2))) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(N_\beta))) & \phi_2(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(N_\beta))) & \dots & \phi_K(\mathbf{X}_{t_m}\mathcal{I}^{\beta}_{t_{m-1}}((N_\beta))) \end{bmatrix},$$

where $\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(1)), \dots, \mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(N_\beta))$ are the states of the paths in bundle $\mathcal{B}_{t_{m-1}}(\beta)$.

- The corresponding vector of option values for these paths

$$\mathbf{V}^{\beta}_{t_m} := \begin{bmatrix} \widehat{V}_{t_m}(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(1))) \\ \widehat{V}_{t_m}(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(2))) \\ \vdots \\ \widehat{V}_{t_m}(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(N_\beta))) \end{bmatrix}.$$

Delta along the paths

- The least squares coefficients computation can be written as:

$$\widehat{\alpha}_{t_m}^{\beta} = (\mathbf{A}_{t_m}^{\beta}{}^{\top}\mathbf{A}_{t_m}^{\beta})^{-1}(\mathbf{A}_{t_m}^{\beta}{}^{\top})\mathbf{V}_{t_m}^{\beta}.$$

- The derivative of the regression coefficients is then given by

$$\begin{aligned}\frac{\partial \alpha_{t_m}^{\beta}}{\partial \mathbf{X}_{t_{m-1}}} &= \frac{\partial (\mathbf{A}_{t_m}^{\beta}{}^{\top}\mathbf{A}_{t_m}^{\beta})^{-1}}{\partial \mathbf{X}_{t_{m-1}}}(\mathbf{A}_{t_m}^{\beta}{}^{\top})\mathbf{V}_{t_m}^{\beta} \\ &+ (\mathbf{A}_{t_m}^{\beta}{}^{\top}\mathbf{A}_{t_m}^{\beta})^{-1}\frac{\partial \mathbf{A}_{t_m}^{\beta}{}^{\top}}{\partial \mathbf{X}_{t_{m-1}}}\mathbf{V}_{t_m}^{\beta} \\ &+ (\mathbf{A}_{t_m}^{\beta}{}^{\top}\mathbf{A}_{t_m}^{\beta})^{-1}(\mathbf{A}_{t_m}^{\beta}{}^{\top})\frac{\partial \mathbf{V}_{t_m}^{\beta}}{\partial \mathbf{X}_{t_{m-1}}},\end{aligned}$$

- The derivative of the matrix inverse can be further expanded as

$$\frac{\partial (\mathbf{A}_{t_m}^{\beta}{}^{\top}\mathbf{A}_{t_m}^{\beta})^{-1}}{\partial \mathbf{X}_{t_{m-1}}} = -(\mathbf{A}_{t_m}^{\beta}{}^{\top}\mathbf{A}_{t_m}^{\beta})^{-1}\left(\frac{\partial \mathbf{A}_{t_m}^{\beta}{}^{\top}}{\partial \mathbf{X}_{t_{m-1}}}\mathbf{A}_{t_m}^{\beta} + \mathbf{A}_{t_m}^{\beta}{}^{\top}\frac{\partial \mathbf{A}_{t_m}^{\beta}}{\partial \mathbf{X}_{t_{m-1}}}\right)(\mathbf{A}_{t_m}^{\beta}{}^{\top}\mathbf{A}_{t_m}^{\beta})^{-1}$$

Delta along the paths

- So, to compute $\frac{\partial \alpha^{\beta}_{t_m}}{\partial \mathbf{X}_{t_{m-1}}}$, we need the quantities $\frac{\partial \mathbf{A}^{\beta}_{t_m}}{\partial \mathbf{X}_{t_{m-1}}}$ and $\frac{\partial \mathbf{V}^{\beta}_{t_m}}{\partial \mathbf{X}_{t_{m-1}}}$.
- The derivative of the regression matrix reads

$$\frac{\partial \mathbf{A}^{\beta}_{t_m}}{\partial \mathbf{X}_{t_{m-1}}} = \begin{bmatrix} \frac{\partial \phi_1(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(1)))}{\partial \mathbf{X}_{t_m}} \frac{\partial \mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(1))}{\partial \mathbf{X}_{t_{m-1}}} & \cdots & \frac{\partial \phi_K(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}1)))}{\partial \mathbf{X}_{t_m}} \frac{\partial \mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(1))}{\partial \mathbf{X}_{t_{m-1}}} \\ \frac{\partial \phi_1(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(2)))}{\partial \mathbf{X}_{t_m}} \frac{\partial \mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(2))}{\partial \mathbf{X}_{t_{m-1}}} & \cdots & \frac{\partial \phi_K(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(2)))}{\partial \mathbf{X}_{t_m}} \frac{\partial \mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(2))}{\partial \mathbf{X}_{t_{m-1}}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \phi_1(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(N_\beta)))}{\partial \mathbf{X}_{t_m}} \frac{\partial \mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(N_\beta))}{\partial \mathbf{X}_{t_{m-1}}} & \cdots & \frac{\partial \phi_K(\mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(N_\beta)))}{\partial \mathbf{X}_{t_m}} \frac{\partial \mathbf{X}_{t_m}(\mathcal{I}^{\beta}_{t_{m-1}}(N_\beta))}{\partial \mathbf{X}_{t_{m-1}}} \end{bmatrix},$$

  where $\frac{\partial \mathbf{X}_{t_m}}{\partial \mathbf{X}_{t_{m-1}}}$ is obtained using the discretization scheme.
- Finally, $\frac{\partial \mathbf{V}^{\beta}_{t_m}}{\partial \mathbf{X}_{t_{m-1}}}$ is given by

$$\frac{\partial \mathbf{V}^{\beta}_{t_m}}{\partial \mathbf{X}_{t_{m-1}}} = \frac{\partial \mathbf{V}^{\beta}_{t_m}}{\partial \mathbf{X}_{t_m}} \frac{\partial \mathbf{X}_{t_m}}{\partial \mathbf{X}_{t_{m-1}}}.$$

Model parameter sensitivities

- The sensitivity of the option value at t_{m-1} w.r.t θ is given by

$$\begin{aligned}\frac{\partial}{\partial\theta}V_{t_{m-1}}(\mathbf{X}_{t_{m-1}}) &= \left(\frac{\partial}{\partial\theta}h\left(\mathbf{X}_{t_{m-1}}\right)\right)\mathbb{1}_{Q_{t_{m-1}}<h(\mathbf{X}_{t_{m-1}})} \\ &+ \left(\frac{\partial}{\partial\theta}Q_{t_{m-1}}\left(\mathbf{X}_{t_{m-1}}\right)\right)\mathbb{1}_{Q_{t_{m-1}}\geq h(\mathbf{X}_{t_{m-1}})}.\end{aligned}$$

- Again, the payoff term is usually trivial to compute.
- The derivative of the $Q_{t_{m-1}}$ for $\mathbf{X}_{t_{m-1}}(n)$ in bundle $\mathcal{B}_{t_{m-1}}(\beta)$ is

$$\begin{aligned}\frac{\partial}{\partial\theta}\widehat{Q}_{t_{m-1}}(\mathbf{X}_{t_{m-1}}(n)) &= \frac{\partial}{\partial\theta}\left(\sum_{k=1}^{K}\widehat{\alpha}^{\beta}_{t_m}(k)\mathbb{E}\left[\phi_k(\mathbf{X}_{t_m})\mid\mathbf{X}_{t_{m-1}}(n)\right]\right) \\ &= \sum_{k=1}^{K}\left(\left(\frac{\partial}{\partial\theta}\widehat{\alpha}^{\beta}_{t_m}(k)\right)\mathbb{E}\left[\phi_k(\mathbf{X}_{t_m})\mid\mathbf{X}_{t_{m-1}}(n)\right]\right. \\ &+ \left.\widehat{\alpha}^{\beta}_{t_m}(k)\frac{\partial}{\partial\theta}\mathbb{E}\left[\phi_k(\mathbf{X}_{t_m})\mid\mathbf{X}_{t_{m-1}}(n)\right]\right)\end{aligned}$$

Model parameter sensitivities

- $\frac{\partial}{\partial\theta}\mathbb{E}\left[\phi_k(\mathbf{X}_{t_m}) \mid \mathbf{X}_{t_{m-1}}(n)\right]$ is usually trivial to compute.
- Following the same idea as before, we can write

$$\begin{aligned}
\frac{\partial \alpha^{\beta}_{t_m}}{\partial\theta} &= \frac{\partial({\mathbf{A}^{\beta}_{t_m}}^{\top}\mathbf{A}^{\beta}_{t_m})^{-1}}{\partial\theta}({\mathbf{A}^{\beta}_{t_m}}^{\top})\mathbf{V}^{\beta}_{t_m} \\
&+ ({\mathbf{A}^{\beta}_{t_m}}^{\top}\mathbf{A}^{\beta}_{t_m})^{-1}\frac{\partial {\mathbf{A}^{\beta}_{t_m}}^{\top}}{\partial\theta}\mathbf{V}^{\beta}_{t_m} \\
&+ ({\mathbf{A}^{\beta}_{t_m}}^{\top}\mathbf{A}^{\beta}_{t_m})^{-1}({\mathbf{A}^{\beta}_{t_m}}^{\top})\frac{\partial \mathbf{V}^{\beta}_{t_m}}{\partial\theta},
\end{aligned}$$

- Similarly, we further expand the inverse derivative as

$$\frac{\partial({\mathbf{A}^{\beta}_{t_m}}^{\top}\mathbf{A}^{\beta}_{t_m})^{-1}}{\partial\theta} = -({\mathbf{A}^{\beta}_{t_m}}^{\top}\mathbf{A}^{\beta}_{t_m})^{-1}\left(\frac{\partial {\mathbf{A}^{\beta}_{t_m}}^{\top}}{\partial\theta}\mathbf{A}^{\beta}_{t_m} + {\mathbf{A}^{\beta}_{t_m}}^{\top}\frac{\partial \mathbf{A}^{\beta}_{t_m}}{\partial\theta}\right)({\mathbf{A}^{\beta}_{t_m}}^{\top}\mathbf{A}^{\beta}_{t_m})^{-1}.$$

Model parameter sensitivities

- We now need the quantities $\frac{\partial \mathbf{A}^\beta_{t_m}}{\partial \theta}$ and $\frac{\partial \mathbf{V}^\beta_{t_m}}{\partial \theta}$.
- The derivative w.r.t parameter θ of the regression matrix is

$$\frac{\partial \mathbf{A}^\beta_{t_m}}{\partial \theta} = \begin{bmatrix} \frac{\partial \phi_1(\mathbf{X}_{t_m}(\mathcal{I}^\beta_{t_{m-1}}(1)))}{\partial \mathbf{X}_{t_m}} \frac{\partial \mathbf{X}_{t_m}(\mathcal{I}^\beta_{t_{m-1}}(1))}{\partial \theta} & \dots & \frac{\partial \phi_K(\mathbf{X}_{t_m}(\mathcal{I}^\beta_{t_{m-1}}1)))}{\partial \mathbf{X}_{t_m}} \frac{\partial \mathbf{X}_{t_m}(\mathcal{I}^\beta_{t_{m-1}}(1))}{\partial \theta} \\ \frac{\partial \phi_1(\mathbf{X}_{t_m}(\mathcal{I}^\beta_{t_{m-1}}(2)))}{\partial \mathbf{X}_{t_m}} \frac{\partial \mathbf{X}_{t_m}(\mathcal{I}^\beta_{t_{m-1}}(2))}{\partial \theta} & \dots & \frac{\partial \phi_K(\mathbf{X}_{t_m}(\mathcal{I}^\beta_{t_{m-1}}(2)))}{\partial \mathbf{X}_{t_m}} \frac{\partial \mathbf{X}_{t_m}(\mathcal{I}^\beta_{t_{m-1}}(2))}{\partial \theta} \\ \vdots & \ddots & \vdots \\ \frac{\partial \phi_1(\mathbf{X}_{t_m}(\mathcal{I}^\beta_{t_{m-1}}(N_\beta)))}{\partial \mathbf{X}_{t_m}} \frac{\partial \mathbf{X}_{t_m}(\mathcal{I}^\beta_{t_{m-1}}(N_\beta))}{\partial \theta} & \dots & \frac{\partial \phi_K(\mathbf{X}_{t_m}(\mathcal{I}^\beta_{t_{m-1}}(N_\beta)))}{\partial \mathbf{X}_{t_m}} \frac{\partial \mathbf{X}_{t_m}(\mathcal{I}^\beta_{t_{m-1}}(N_\beta))}{\partial \theta} \end{bmatrix},$$

where $\frac{\partial \mathbf{X}_{t_m}}{\partial \theta}$ is usually easy to obtain from the discretization scheme.

- Since $\mathbf{V}^\beta_{t_m} := \mathbf{V}^\beta_{t_m}(\mathbf{X}_{t_m}, \theta)$, the derivative of the option price vector is

$$\frac{\partial \mathbf{V}^\beta_{t_m}}{\partial \theta} | \mathcal{F}_{t_{m-1}} = \frac{\partial \mathbf{V}^\beta_{t_m}}{\partial \mathbf{X}_{t_m}} \frac{\partial \mathbf{X}_{t_m}}{\partial \theta} + \frac{\partial \mathbf{V}^\beta_{t_m}}{\partial \theta},$$

where $\frac{\partial \mathbf{V}^\beta_{t_m}}{\partial \mathbf{X}_{t_m}}$ is exactly the **Delta** sensitivity.

Adjoint formulation

- Often monitoring dates are different from discretization time points.
- For instance, when an exact discretization is not available, intermediate simulated times need to be introduced to preserve (increase) the accuracy.
- When this occurs, the derivatives need to be propagated along the intermediate steps.
- For that, the computation of the path-level sensitivities with SGBM admits the adjoint formulation, as described in Giles and Glasserman.
- With the difference that the recursion takes place between t_m, and t_{m-1} in *Rolling Adjoints* rather than between t_M, and t_0, in *Smoking Adjoints*.
- The adjoint mode can provide a significant gain in the computational cost when the number of inputs is large.

Adjoint formulation - Delta sensitivity

- Let $t_{m-1} = t_{m_0}, \ldots, t_{m_l}, \ldots, t_{m_L} = t_m$ denote the sub-discretization between t_{m-1} and t_m, and

$$\Delta_{m_l} := \frac{\partial \mathbf{X}_{t_{m_l}}}{\partial \mathbf{X}_{t_{m_{l-1}}}} = \frac{\partial F_{m_{l-1}}(\mathbf{X}_{t_{m_{l-1}}}, \mathbf{Z}_{t_{m_{l-1}}}, \theta)}{\partial \mathbf{X}_{t_{m_{l-1}}}}.$$

- $\frac{\partial \phi_k(\mathbf{X}_{t_{m_L}})}{\partial \mathbf{X}_{t_{m_L}}} \frac{\partial \mathbf{X}_{t_{m_L}}}{\partial \mathbf{X}_{t_{m_0}}}$, $\frac{\partial \mathbf{V}^{\beta}_{t_{m_L}}}{\partial \mathbf{X}_{t_{m_L}}} \frac{\partial \mathbf{X}_{t_{m_L}}}{\partial \mathbf{X}_{t_{m_0}}}$ are computed using the recursion

$$\frac{\partial \phi_k(\mathbf{X}_{t_{m_L}})}{\partial \mathbf{X}_{t_{m_L}}} \Delta_{m_L} \Delta_{m_{L-1}} \ldots \Delta_{m_0}, \qquad \frac{\partial \mathbf{V}^{\beta}_{t_{m_L}}}{\partial \mathbf{X}_{t_{m_L}}} \Delta_{m_L} \Delta_{m_{L-1}} \ldots \Delta_{m_0}.$$

- Adjoint (from left to right) vs. forward (from right to left),

$$\overrightarrow{\text{Adjoint}} \qquad \overleftarrow{\text{Forward}}$$

$$\frac{\partial \phi_k(\mathbf{X}_{t_{m_L}})}{\partial \mathbf{X}_{t_{m_L}}} \Delta_{m_L} \Delta_{m_{L-1}} \ldots \Delta_{m_0} \qquad \frac{\partial \phi_k(\mathbf{X}_{t_{m_L}})}{\partial \mathbf{X}_{t_{m_L}}} \Delta_{m_L} \Delta_{m_{L-1}} \ldots \Delta_{m_0}$$

$$\frac{\partial \mathbf{V}^{\beta}_{t_{m_L}}}{\partial \mathbf{X}_{t_{m_L}}} \Delta_{m_L} \Delta_{m_{L-1}} \ldots \Delta_{m_0}, \qquad \frac{\partial \mathbf{V}^{\beta}_{t_{m_L}}}{\partial \mathbf{X}_{t_{m_L}}} \Delta_{m_L} \Delta_{m_{L-1}} \ldots \Delta_{m_0}.$$

Adjoint formulation - Model parameter sensitivities

- We need to compute

$$\frac{\partial \phi_k(\mathbf{X}_{t_m})}{\partial \mathbf{X}_{t_m}} \frac{\partial \mathbf{X}_{t_m}}{\partial \theta}, k = 1 \ldots, K.$$

- Denoting

$$\Theta_{m_l} := \frac{\partial \mathbf{X}_{t_{m_l}}}{\partial \theta} = \frac{\partial}{\partial \theta} F_{m_{l-1}}(\mathbf{X}_{t_{m_{l-1}}}, \mathbf{Z}_{t_{m_{l-1}}}, \theta).$$

- $\frac{\partial \mathbf{X}_{t_{m_L}}}{\partial \theta}$ is recursively calculated using the chain rule as

$$\Theta_{m_l} = \frac{\partial F_{m_{l-1}}(\mathbf{X}_{t_{m_{l-1}}}, \mathbf{Z}_{t_{m_{l-1}}}, \theta)}{\partial \mathbf{X}_{t_{m_{l-1}}}} \Theta_{m_{l-1}} + \frac{\partial F_{m_{l-1}}(\mathbf{X}_{t_{m_{l-1}}}, \mathbf{Z}_{t_{m_{l-1}}}, \theta)}{\partial \theta},$$

where $l = 1, \ldots, L$, with initial condition $\Theta_{m_0} = 0$.

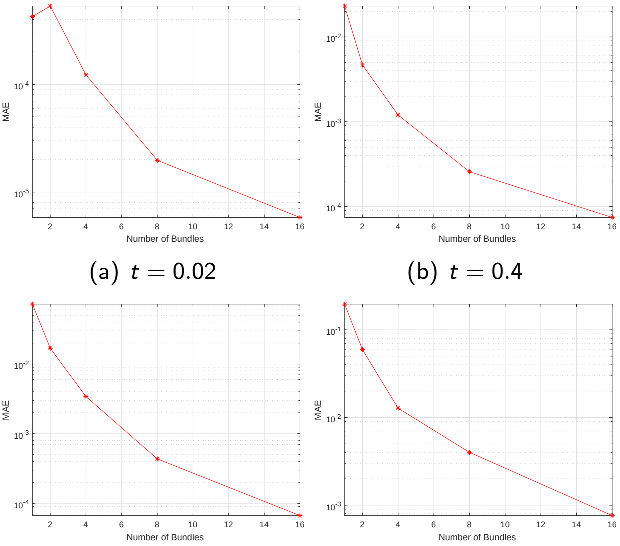
- This recursion admits again both the forward formulation and the adjoint formulation.

Numerical results

- Geometric Brownian Motion, $90,000$ paths.
- European, Bermudan and spread options. Two sets:

Set I	$\mathbf{X}_{t_0}$	36, 40, 44
	σ	10%, 20%,40%
	r	0.06
	Strike K	40
	M	50
	T	1 year
Set II	$\mathbf{X}_{t_0} := \{S^1_{t_0}, S^2_{t_0}\}$	[100, 100]
	$\sigma := \{\sigma^1, \sigma^2\}$	[15% 15%]
	r	0.03
	Strike K	5
	M	8
	ρ_{12}	0.5
	T	1 year

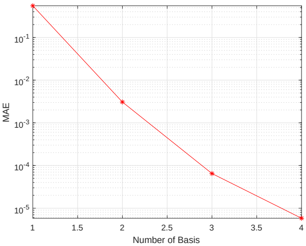
European option - Delta convergence in bundles



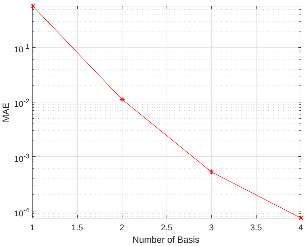
(a) $t = 0.02$ (b) $t = 0.4$

(c) $t = 0.7$ (d) $t = 0.98$

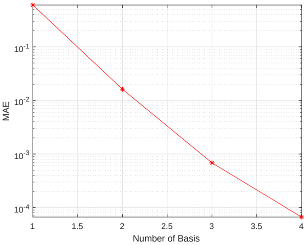
European option - Delta convergence in basis functions



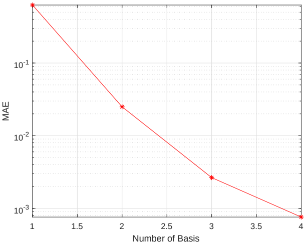
(a) $t = 0.02$



(b) $t = 0.4$

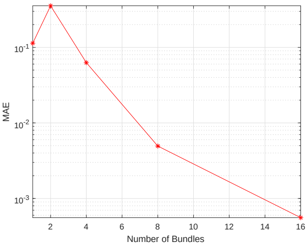


(c) $t = 0.7$

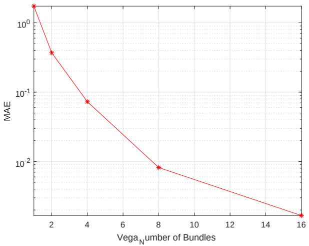


(d) $t = 0.98$

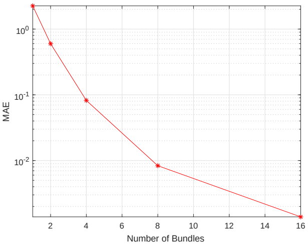
European option - Vega convergence in bundles



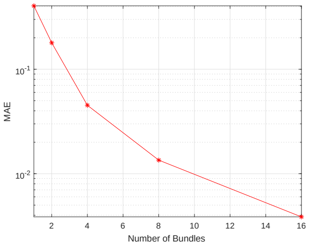
(a) $t = 0.02$



(b) $t = 0.4$

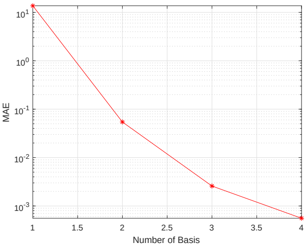


(c) $t = 0.7$

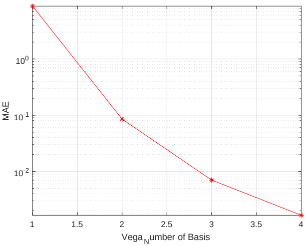


(d) $t = 0.98$

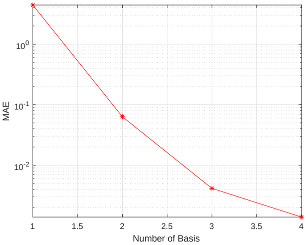
European option - Vega convergence in basis functions



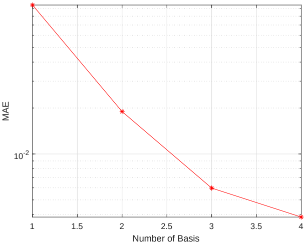
(a) $t = 0.02$



(b) $t = 0.4$



(c) $t = 0.7$



(d) $t = 0.98$

Bermudan option - Greeks at t_0

$\mathbf{X}_{t_0}$	COS Delta	SGBM Delta (s.e.)	Error SGBM	LSMC1 Delta (s.e.)	Error LSMC1	LSMC2 Delta (s.e.)	Error LSMC2
36	-0.695	-0.695 (0.6e-5)	-0.0001	-0.711 (0.0213)	0.0159	-0.972 (0.227)	-0.2770
40	-0.404	-0.404 (0.5e-5)	0.0003	-0.402 (0.0190)	-0.0019	-0.463 (0.033)	-0.0591
44	-0.213	-0.214 (0.9e-5)	0.0009	-0.227 (0.0080)	0.0141	-0.253 (0.031)	-0.0396

Table: t_0 Delta values for Bermudan put option on a single asset for different initial asset prices. The values in brackets are the standard errors from thirty trials.

$\mathbf{X}_{t_0}$	COS Vega	SGBM Vega (s.e.)	Error SGBM	LSMC1 Vega (s.e.)	Error LSMC1	LSMC2 Vega (s.e.)	Error LSMC2
36	10.955	10.920 (0.001)	-0.0348	11.099 (0.070)	0.1445	10.734 (0.231)	-0.2209
40	14.747	14.752 (0.001)	0.0049	14.890 (0.099)	0.1438	14.730 (0.057)	-0.0170
44	12.524	12.616 (0.003)	0.0924	12.556 (0.062)	0.0318	12.536 (0.051)	0.0126

Table: t_0 Vega values for Bermudan put option on a single asset for different initial asset prices. The values in brackets are the standard errors from thirty trials.

Bermudan option - Greeks at t_0

σ	COS Vega	SGBM Vega (s.e.)	Error SGBM	LSMC1 Vega (s.e.)	Error LSMC1	LSMC2 Vega (s.e.)	Error LSMC2
10%	13.360	13.402 (0.002)	0.0416	13.526 (0.062)	0.1652	13.285 (0.066)	-0.0754
20%	14.747	14.750 (0.001)	0.0034	14.931 (0.084)	0.1841	14.730 (0.057)	-0.0170
40%	15.055	15.053 (0.002)	-0.0019	15.188 (0.104)	0.1336	15.115 (0.087)	0.0598

Table: t_0 Vega values for Bermudan put option on a single asset for different asset volatilities. The initial asset value is $\mathbf{X}_{t_0} = 40$.

$\mathbf{X}_{t_0}$	COS Vega	SGBM Vega (s.e.)	Error SGBM	LSMC1 Vega (s.e.)	Error LSMC1	LSMC2 Vega (s.e.)	Error LSMC2
34.5	6.794	6.757 (0.0008)	-0.0372	7.062 (0.212)	0.2677	6.866 (0.433)	0.0719
35	8.383	8.342 (0.001)	0.0414	8.621 (0.119)	0.2374	8.076 (0.149)	-0.3075
35.5	9.771	9.731 (0.001)	0.0397	10.224 (0.103)	0.4529	9.450 (0.161)	-0.3206

Table: t_0 Vega values for Bermudan put option on a single asset for a case where the initial asset price is close to the early-exercise boundary, $\mathbf{X}_{t_0} = 34.5$.

Bermudan spread option - Greeks at t_0

	SGBM extended Delta (s.e)	SGBM BR Delta (s.e)	LSMC1 BR Delta (s.e.)	LSMC2 BR Delta (s.e.)
$\frac{\partial V_{t_0}}{\partial S^1_{t_0}}$	0.4020 (0.2e-4)	0.4021 (0.1e-3)	0.4029 (0.011)	0.4570 (0.083)
$\frac{\partial V_{t_0}}{\partial S^2_{t_0}}$	-0.3448 (0.2e-4)	-0.3453 (0.1e-3)	-0.3446 (0.010)	-0.3795 (0.085)

Table: t_0 Delta values for Bermudan spread option on two assets.

	SGBM extended Vega (s.e)	SGBM BR Vega (s.e)	LSMC1 BR Vega (s.e.)	LSMC2 BR Vega (s.e.)
$\frac{\partial V_{t_0}}{\partial \sigma_1}$	20.6082 (0.016)	20.7551 (0.025)	20.4900 (0.124)	20.5136 (0.198)
$\frac{\partial V_{t_0}}{\partial \sigma_2}$	16.8822 (0.013)	17.0611 (0.017)	17.0022 (0.089)	17.1409 (0.155)

Table: Vega t_0 values for Bermudan spread option on two assets.

Case	SGBM extended	SGBM BR	LSMC1 BR	LSMC2 BR
Single Asset (50 monitoring dates)	4.5s	10s	2s	4.2s
Two Asset (8 monitoring dates)	3s	12s	4s	7s

Table: The computational time of 30 trials.

Conclusions

- We have presented an approach to compute sensitivities w.r.t state space and model parameters along the path for early-exercise options.
- The approach is applicable to regress-later schemes like SGBM.
- Through the examples we numerically illustrate study the convergence of the method and demonstrate the stability of the method.
- The sensitivities along the paths are computed without significant computational and memory overhead.
- Future work:
  - Compute MVA for SIMM based initial margins.
  - Sensitivities in energy market complex options.

References

Shashi Jain, Álvaro Leitao, and Cornelis W. Oosterlee.

Rolling adjoints: fast Greeks along Monte Carlo scenarios for early-exercise options, 2017.

Submitted to Applied Mathematics and Computation. Available at SSRN: https://ssrn.com/abstract=3093846.

Shashi Jain and Cornelis W. Oosterlee.

The Stochastic Grid Bundling Method: Efficient pricing of Bermudan options and their Greeks.

Applied Mathematics and Computation, 269:412–431, 2015.

Acknowledgements & Questions

EXCELENCIA
MARÍA
DE MAEZTU

Thanks to support from MDM-2014-0445

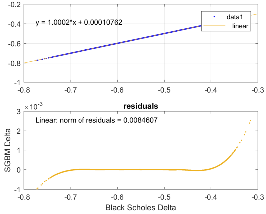
More: `leitao@ub.edu` and `alvaroleitao.github.io`

Thank you for your attention

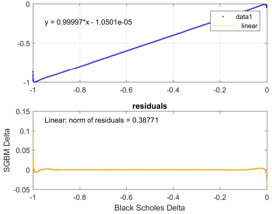
Bonus

- Residual errors.
- Basic European option experiment.
- Influence of bundling on regress-later approaches.

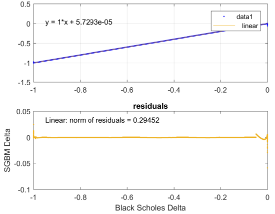
European option - Deltas along the paths



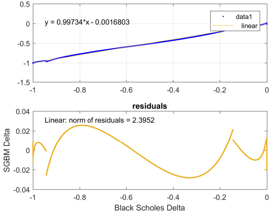
(a) $t = 0.02$



(b) $t = 0.4$

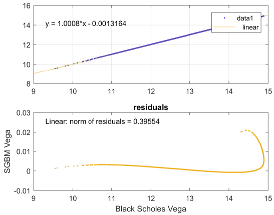


(c) $t = 0.7$

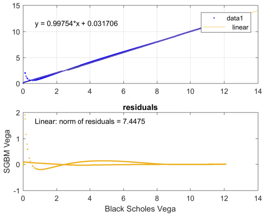


(d) $t = 0.98$

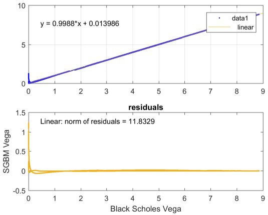
European option - Vega along the paths



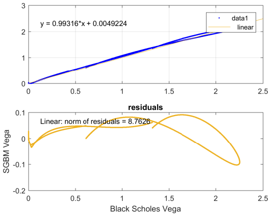
(e) $t = 0.02$



(f) $t = 0.4$



(g) $t = 0.7$



(h) $t = 0.98$

European option - Greeks at t_0

$\mathbf{X}_{t_0}$	SGBM Delta (s.e.)	BS Delta	error	SGBM Vega (s.e.)	BS Vega	error	SGBM Gamma (s.e.)	BS Gamma	error
36	-0.5504 (0.2e-5)	-0.5504	0.1e-4	14.2526 (0.0005)	14.2469	0.0057	0.0550 (0.2e-5)	0.0550	-0.03e-4
40	-0.3445 (0.1e-5)	-0.3445	0.1e-4	14.7399 (0.0006)	14.7308	0.0091	0.0460 (0.2e-5)	0.0460	-0.06e-4
44	-0.1903 (0.4e-5)	-0.1903	0.4e-4	11.9702 (0.0007)	11.9542	0.0160	0.0309 (0.2e-5)	0.0309	-0.12e-4

Table: The t_0 Delta, Vega, and Gamma values computed using SGBM. The values in brackets are corresponding standard errors for SGBM for 30 trials.

Bundling on regress-later approaches

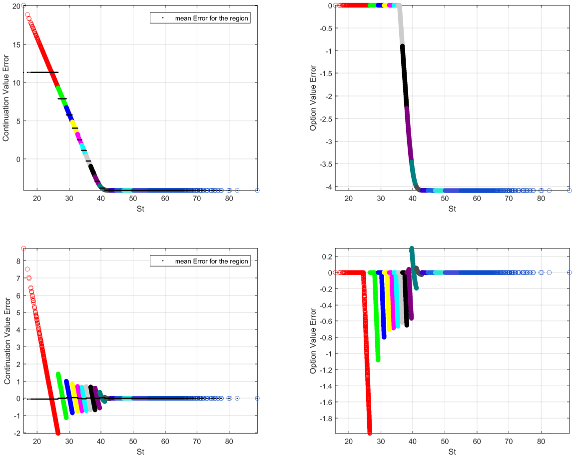


Figure: Basic regress-later vs. SGBM. Local errors in continuation and option values at t_{M-1}. The different colors indicate different regions. $K = 1$.

Bundling on regress-later approaches

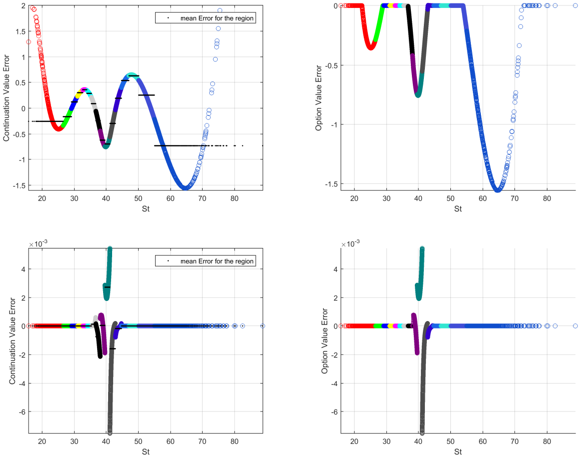


Figure: Basic regress-later vs. SGBM. Local errors in continuation and option values at t_{M-1}. The different colors indicate different regions. $K = 6$.