AD-on-LSMC for MVA and CVA Greeks: Simplifications and Efficiencies

Andrew McClelland and Serguei Issakov
Quantitative Research, Numerix

Alexander Antonov
Quantitative Research, Standard Chartered

QuantMinds International, Lisbon

May 17, 2018

Presentation Outline

- CVA Greeks and MVA via "Future" Greeks
- Future Greeks as a by-product of AD-on-LSMC
- AD efficiencies for LSMC: large-sample regression coefficient dependencies

Presentation Outline

- CVA Greeks and MVA via “Future” Greeks
- Future Greeks as a by-product of AD-on-LSMC
- AD efficiencies for LSMC: large-sample regression coefficient dependencies

Future Greeks for CVA Greeks and MVA

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0 \left[\int_0^T e^{-R(t)} (V(t))^+ \lambda(t) \, dt \right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \partial_\theta \, \mathbb{E}_0 \left[\int_0^T e^{-R(t)} (V(t))^+ \lambda(t) \, dt \right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0 \left[\int_0^T e^{-R(t)} \text{IM}(t) \, s(t) \, dt \right]$$

Future Greeks for CVA Greeks and MVA

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T e^{-R(t)}(V(t))^+\,\lambda(t)\,dt\right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \partial_\theta\,\mathbb{E}_0\left[\int_0^T e^{-R(t)}(V(t))^+\,\lambda(t)\,dt\right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0\left[\int_0^T e^{-R(t)}\text{IM}(t)\,s(t)\,dt\right]$$

numerix

Future Greeks for CVA Greeks and MVA

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \partial_\theta \, \mathbb{E}_0\left[\int_0^T e^{-R(t)}(V(t))^+ \, \lambda(t) \, dt\right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0\left[\int_0^T e^{-R(t)} \text{IM}(t) \, s(t) \, dt\right]$$

Future Greeks for CVA Greeks and MVA

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\mathrm{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \mathrm{CVA} = \partial_\theta \, \mathbb{E}_0\left[\int_0^T e^{-R(t)} (V(t))^+ \, \lambda(t) \, dt\right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\mathrm{MVA} = \mathbb{E}_0\left[\int_0^T e^{-R(t)} \mathrm{IM}(t) \, s(t) \, dt\right]$$

Future Greeks for CVA Greeks and MVA

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\mathrm{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \mathrm{CVA} = \partial_\theta \, \mathbb{E}_0\left[\int_0^T e^{-R(t)} (V(t))^+ \, \lambda(t) \, dt\right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\mathrm{MVA} = \mathbb{E}_0\left[\int_0^T e^{-R(t)} \mathrm{IM}(t) \, s(t) \, dt\right]$$

Future Greeks for CVA Greeks and MVA

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\mathrm{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \mathrm{CVA} = \partial_\theta \, \mathbb{E}_0\left[\int_0^T {\color{red}(V(t))^+} \, dt\right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\mathrm{MVA} = \mathbb{E}_0\left[\int_0^T e^{-R(t)}\mathrm{IM}(t)\, s(t)\, dt\right]$$

numerix

Future Greeks for CVA Greeks and MVA

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\mathrm{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \mathrm{CVA} = \partial_\theta \, \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\mathrm{MVA} = \mathbb{E}_0\left[\int_0^T e^{-R(t)} \mathrm{IM}(t) \, s(t) \, dt\right]$$

Future Greeks for CVA Greeks and MVA

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \partial_\theta \, \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0\left[\int_0^T e^{-R(t)} \text{IM}(t) \, s(t) \, dt\right]$$

Future Greeks for CVA Greeks and MVA

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\mathrm{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \mathrm{CVA} = \mathbb{E}_0\left[\int_0^T \partial_\theta (V(t))^+ \, dt\right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\mathrm{MVA} = \mathbb{E}_0\left[\int_0^T e^{-R(t)} \mathrm{IM}(t)\, s(t)\, dt\right]$$

Future Greeks for CVA Greeks and MVA

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt\right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0\left[\int_0^T e^{-R(t)} \text{IM}(t) \, s(t) \, dt\right]$$

Future Greeks for CVA Greeks and MVA

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt\right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0\left[\int_0^T e^{-R(t)} \text{IM}(t) \, s(t) \, dt\right]$$

Future Greeks for CVA Greeks and MVA

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt\right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0\left[\int_0^T e^{-R(t)} \text{IM}(t) \, s(t) \, dt\right]$$

Future Greeks for CVA Greeks and MVA

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0 \left[\int_0^T (V(t))^+ \, dt \right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \mathbb{E}_0 \left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt \right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0 \left[\int_0^T \text{IM}(t) \, dt \right]$$

Future Greeks for CVA Greeks and MVA

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt\right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0\left[\int_0^T \text{IM}(\partial_{Q(t)} V(t)) \, dt\right]$$

Future Greeks for CVA Greeks and MVA

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)}\, \partial_\theta V(t)\, dt\right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0\left[\int_0^T \text{IM}(\partial_\theta V(t), \partial_\theta Q(t))\, dt\right]$$

Presentation Outline

- CVA Greeks and MVA via "Future" Greeks
- Future Greeks as a by-product of AD-on-LSMC
- AD efficiencies for LSMC: large-sample regression coefficient dependencies

Presentation Outline

- CVA Greeks and MVA via "Future" Greeks
- Future Greeks as a by-product of AD-on-LSMC
- AD efficiencies for LSMC: large-sample regression coefficient dependencies

Future Greeks by AD in the LSMC Context 1

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+\, dt\right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)}\, \partial_\theta V(t)\, dt\right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0\left[\int_0^T \text{IM}(\partial_\theta V(t), \partial_\theta Q(t))\, dt\right]$$

Future Greeks by AD in the LSMC Context 1

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt\right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0\left[\int_0^T \text{IM}(\partial_\theta V(t), \partial_\theta Q(t)) \, dt\right]$$

numerix

Future Greeks by AD in the LSMC Context 1

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} \approx \frac{1}{N_P} \sum_{p=1}^{N_P} \sum_{i=1}^{N_T} (V_{p,i})^+ \Delta_i$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \mathbb{E}_0 \left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt \right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0 \left[\int_0^T \text{IM}(\partial_\theta V(t), \partial_\theta Q(t)) \, dt \right]$$

Future Greeks by AD in the LSMC Context 1

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} \approx \frac{1}{N_P} \sum_{p=1}^{N_P} \sum_{i=1}^{N_T} (V_{p,i})^+ \, \Delta_i$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \mathbb{E}_0 \left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt \right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0 \left[\int_0^T \text{IM}(\partial_\theta V(t), \partial_\theta Q(t)) \, dt \right]$$

numerix

Future Greeks by AD in the LSMC Context 1

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} \approx \frac{1}{N_P} \sum_{p=1}^{N_P} \sum_{i=1}^{N_T} (V_{p,i})^+ \Delta_i$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \mathbb{E}_0 \left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt \right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0 \left[\int_0^T \text{IM}(\partial_\theta V(t), \partial_\theta Q(t)) \, dt \right]$$

Future Greeks by AD in the LSMC Context 1

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\mathrm{CVA} \approx \frac{1}{N_P} \sum_{p=1}^{N_P} \sum_{i=1}^{N_T} (\hat{V}_{p,i})^+ \Delta_i$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \mathrm{CVA} = \mathbb{E}_0 \left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt \right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\mathrm{MVA} = \mathbb{E}_0 \left[\int_0^T \mathrm{IM}(\partial_\theta V(t), \partial_\theta Q(t)) \, dt \right]$$

Future Greeks by AD in the LSMC Context 1

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} \approx \frac{1}{N_P} \sum_{p=1}^{N_P} \sum_{i=1}^{N_T} (\hat{V}_{p,i})^+ \, \Delta_i$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \mathbb{E}_0 \left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt \right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0 \left[\int_0^T \text{IM}(\partial_\theta V(t), \partial_\theta Q(t)) \, dt \right]$$

Future Greeks by AD in the LSMC Context 1

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} \approx \frac{1}{N_P} \sum_{p=1}^{N_P} \sum_{i=1}^{N_T} (\hat{V}_{p,i})^+ \Delta_i$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} = \mathbb{E}_0 \left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt \right]$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0 \left[\int_0^T \text{IM}(\partial_\theta V(t), \partial_\theta Q(t)) \, dt \right]$$

Future Greeks by AD in the LSMC Context 1

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} \approx \frac{1}{N_P} \sum_{p=1}^{N_P} \sum_{i=1}^{N_T} (\hat{V}_{p,i})^+ \, \Delta_i$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} \approx \frac{1}{N_P} \sum_{p=1}^{N_P} \sum_{i=1}^{N_T} 1_{(\hat{V}_{p,i}>0)} \partial_\theta \hat{V}_{p,i} \, \Delta_i$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} \approx \frac{1}{N_P} \sum_{p=1}^{N_P} \sum_{i=1}^{N_T} \text{IM}(\partial_\theta \hat{V}_{p,i}, \partial_\theta Q(t)) \, \Delta_i$$

Future Greeks by AD in the LSMC Context 1

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} \approx \frac{1}{N_P} \sum_{p=1}^{N_P} \sum_{i=1}^{N_T} (\hat{V}_{p,i})^+ \, \Delta_i$$

- CVA Greeks computed by chain rule, involves parameter θ sensitivities

$$\partial_\theta \text{CVA} \approx \frac{1}{N_P} \sum_{p=1}^{N_P} \sum_{i=1}^{N_T} 1_{(\hat{V}_{p,i}>0)} \partial_\theta \hat{V}_{p,i} \, \Delta_i$$

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} \approx \frac{1}{N_P} \sum_{p=1}^{N_P} \sum_{i=1}^{N_T} \text{IM}(\partial_\theta \hat{V}_{p,i}, \partial_\theta Q(t)) \, \Delta_i$$

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{p,i} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow V_{p,i} \approx \phi(X_{i,p}) \cdot \beta$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta\hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta\hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{p,i} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow V_{p,i} \approx \phi(X_{i,p}) \cdot \beta$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{p,i} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow V_{p,i} \approx \phi(X_{i,p}) \cdot \beta$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta\hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta\hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{p,i} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow V_{p,i} \approx \phi(X_{i,p}) \cdot \beta$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{p,i} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow V_{p,i} \approx \phi(X_{i,p}) \cdot \beta$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{p,i} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta\hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta\hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{p,i} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta\hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta\hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{p,i} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta\hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta\hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta}(\theta) = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}(\theta)$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

LSMC Computational Graph

Breakdown of LSMC Dependencies

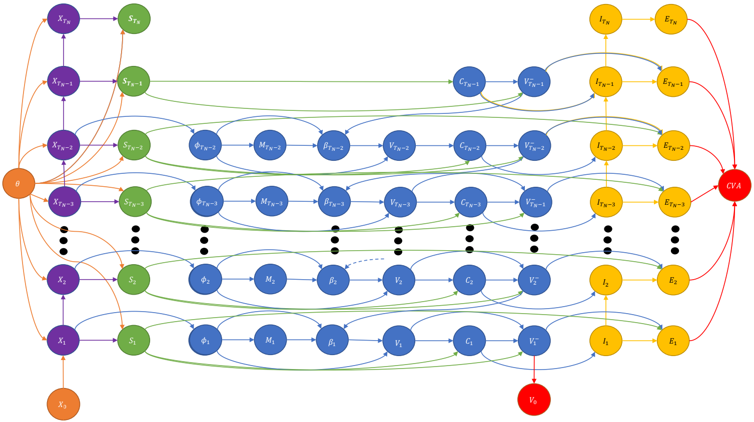

Figure: The LSMC computational graph with dependencies relevant for AD

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta}(\theta) = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}(\theta)$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta}(\theta) = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}(\theta)$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta}(\theta) = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}(\theta)$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \partial_\theta \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \partial_\theta \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Future Greeks by AD in the LSMC Context 2

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Presentation Outline

- CVA Greeks and MVA via "Future" Greeks
- Future Greeks as a by-product of AD-on-LSMC
- AD efficiencies for LSMC: large-sample regression coefficient dependencies

Presentation Outline

- CVA Greeks and MVA via "Future" Greeks
- Future Greeks as a by-product of AD-on-LSMC
- AD efficiencies for LSMC: large-sample regression coefficient dependencies

Coefficient Behavior and Dependencies in Large Samples

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Coefficient Behavior and Dependencies in Large Samples

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta}(\theta) = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}(\theta)$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Coefficient Behavior and Dependencies in Large Samples

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta}(\theta, X_i) = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}(\theta, X_i)$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Coefficient Behavior and Dependencies in Large Samples

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta}(\theta, X_i) = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}(\theta, X_i)$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Coefficient Behavior and Dependencies in Large Samples

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta}(\theta, X_i) = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}(\theta, X_i)$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Coefficient Behavior and Dependencies in Large Samples

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta}(\theta, X_i) = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}(\theta, X_i)$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_{X_i}\hat{\beta} = \partial_{X_i}((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1})$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Coefficient Behavior and Dependencies in Large Samples

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta}(\theta, X_i) = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}(\theta, X_i)$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_{X_i}\hat{\beta} = \partial_{X_i}\left((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}\right)$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Coefficient Behavior and Dependencies in Large Samples

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\lim_{N_P \to \infty} \hat{\beta}(\theta, X_i) = \lim_{N_P \to \infty} (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}(\theta, X_i)$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_{X_i}\hat{\beta} = \partial_{X_i}\left((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}\right)$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Coefficient Behavior and Dependencies in Large Samples

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\lim_{N_P \to \infty} \hat{\beta}(\theta, X_i) = \lim_{N_P \to \infty} (\phi(X_i)'\phi(X_i))^{-1} \phi(X_i)' \hat{V}_{i+1}(\theta, X_i) = \beta_\infty$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_{X_i} \hat{\beta} = \partial_{X_i} \left((\phi(X_i)'\phi(X_i))^{-1} \phi(X_i)' \hat{V}_{i+1} \right)$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Coefficient Behavior and Dependencies in Large Samples

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\lim_{N_P \to \infty} \hat{\beta}(\theta, X_i) = \lim_{N_P \to \infty} (\phi(X_i)' \phi(X_i))^{-1} \phi(X_i)' \hat{V}_{i+1}(\theta, X_i) = \beta_\infty$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\partial_{X_i} \hat{\beta} = \partial_{X_i} \left((\phi(X_i)' \phi(X_i))^{-1} \phi(X_i)' \hat{V}_{i+1} \right)$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Coefficient Behavior and Dependencies in Large Samples

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta}(\theta, X_i) = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}(\theta, X_i) = \beta_\infty$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\lim_{N_P \to \infty} \partial_{X_i} \hat{\beta} = \lim_{N_P \to \infty} \partial_{X_i} \left((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1} \right)$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Coefficient Behavior and Dependencies in Large Samples

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta}(\theta, X_i) = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}(\theta, X_i) = \beta_\infty$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\lim_{N_P\to\infty} \partial_{X_i}\hat{\beta} = \lim_{N_P\to\infty} \partial_{X_i}\left((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}\right) = 0$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

numerix

Coefficient Behavior and Dependencies in Large Samples

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed θ-dependence: $V(t_i, X_{p,i}, \theta) \approx \phi(X_{p,i}) \cdot \beta(\theta)$

$$\hat{\beta}(\theta, X_i) = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}(\theta, X_i) = \beta_\infty$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\lim_{N_P\to\infty} \partial_{X_i}\hat{\beta} = \lim_{N_P\to\infty} \partial_{X_i}\left((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}\right) = 0$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Coefficient Behavior and Dependencies in Large Samples

- LSMC for V_i: regress V_{i+1} onto N_B basis functions $\phi(X_{p,i})$

$$V_{i,p} = \mathbb{E}[V(t_{i+1}, X(t_{i+1}))|X_{p,i}] \longrightarrow \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

- Regression coefficients embed the θ dependence

$$\lim_{N_P \to \infty} \hat{\beta}_i = \lim_{N_P \to \infty} (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1} = \hat{\beta}_\infty$$

- AD: chain rule on recursion & intermediate sensitivities comp'd at run time

$$\lim_{N_P \to \infty} \partial_{X_i}\hat{\beta}\, \partial_\theta X_i = \lim_{N_P \to \infty} \partial_{X_i}\left((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}\right) \partial_\theta X_i = 0$$

- Can evaluate full chain in tangent or adjoint mode
- Good in theory, but how well does $\partial_\theta \hat{V}_{p,i}$ approximate $\partial_\theta V_{p,i}$ in practice?

Presentation Outline

- CVA Greeks and MVA via "Future" Greeks
- Future Greeks as a by-product of AD-on-LSMC
- AD efficiencies for LSMC: large-sample regression coefficient dependencies

Presentation Outline

- CVA Greeks and MVA via “Future” Greeks
- Future Greeks as a by-product of AD-on-LSMC
- AD efficiencies for LSMC: large-sample regression coefficient dependencies

Presentation Outline

- CVA Greeks and MVA via “Future” Greeks
- Future Greeks as a by-product of AD-on-LSMC
- AD efficiencies for LSMC: large-sample regression coefficient dependencies

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \text{CVA} = \partial_\theta \text{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \text{CVA} = \partial_\theta \text{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \text{CVA} = \partial_\theta \text{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \text{CVA} = \partial_\theta \text{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\mathrm{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \mathrm{CVA} = \partial_\theta \mathrm{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \mathrm{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\mathrm{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \mathrm{CVA} = \partial_\theta \mathrm{CVA}\,(\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \mathrm{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)}\, \partial_\theta V(t)\, dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\mathrm{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \mathrm{CVA} = \partial_\theta \mathrm{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \mathrm{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \text{CVA} = \partial_\theta \text{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \text{CVA} = \partial_\theta \text{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta \mathbb{E}_t[C(t,T)] \, dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \text{CVA} = \partial_\theta \text{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \mathbb{E}_t[\partial_\theta C(t,T)] \, dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \text{CVA} = \partial_\theta \text{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T \mathbb{E}_t[1_{(V(t)>0)} \, \partial_\theta C(t,T)] \, dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \text{CVA} = \partial_\theta \text{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \text{CVA} = \int_0^T \mathbb{E}_0[\mathbb{E}_t[1_{(V(t)>0)} \, \partial_\theta C(t, T)]] \, dt$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \text{CVA} = \partial_\theta \text{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \text{CVA} = \int_0^T \mathbb{E}_0[1_{(V(t)>0)} \, \partial_\theta C(t,T)] \, dt$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \text{CVA} = \partial_\theta \text{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \dots, f_{N_F}, \sigma_1, \dots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta C(t,T) \, dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \text{CVA} = \partial_\theta \text{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta V(t) \, dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\mathrm{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \mathrm{CVA} = \partial_\theta \mathrm{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \mathrm{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta \left(V(t) - 1_{(S(t^c)-V(t^c)>0)}(V(t) - S(t))\right) dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \text{CVA} = \partial_\theta \text{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta \left(V(t) - 1_{(S(t^c)-V(t^c)>0)}(V(t) - S(t))\right) dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\mathrm{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \mathrm{CVA} = \partial_\theta \mathrm{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \mathrm{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta \left(V(t) - 1_{(S(t^c) - V(t^c) > 0)} (V(t) - S(t))\right) dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \text{CVA} = \partial_\theta \text{CVA} \, (\partial_\theta Q)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t)$ for callables[1]

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta \left(V(t) - 1_{(S(t^c)-V(t^c)>0)}(V(t) - S(t))\right) dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

CVA Greeks: Usage and Calculation

- CVA is value of credit risk in derivatives portfolio (or hedging cost)

$$\text{CVA} = \mathbb{E}_0\left[\int_0^T (V(t))^+ \, dt\right]$$

- Greeks against quotes, Q, *eg*. swap rates or vols, computed via Jacobians

$$\partial_Q \text{CVA} = \partial_\theta \text{CVA} \left(\partial_\theta Q\right)^{-1}$$

- θ is a parameter vector, possibly including initial states, X_0, *eg*. FX spot
- HW-1F *eg*. has forward rate & vol knots, $\theta = [f_1, \ldots, f_{N_F}, \sigma_1, \ldots, \sigma_{N_\sigma}]$
- There is a *formal* requirement for $\partial_\theta V(t^{c_1}), \ldots, V(t^{c_{N_C}})$ for callables[1]

$$\partial_\theta \text{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \, \partial_\theta \left(V(t) - 1_{(S(t^c) - V(t^c) > 0)}(V(t) - S(t))\right) dt\right]$$

[1] Conditioning: Andreasen (14), Indicators: Antonov *et al* ('16) & Capriotti *et al* ('16)

MVA: Motivation and Logistics 1

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR

$$\text{MVA} = \mathbb{E}_0 \left[\int_0^T \text{IM}(\partial_{Q(t)} V(t))\, dt \right]$$

- IM is additional collateral to mitigate counterparty risk over MPoR ($\sim$ 10D)
- Bilateral IM: both c/parties post to 3^{rd}-party custodians $\implies$ needs funding
- In practice, portfolio hedges attract bilateral &/or clearing-house IM too
- MVA reflects funding costs in valuations $\implies$ spectre of FVA debate

MVA: Motivation and Logistics 1

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR[2]

$$\text{MVA} = \mathbb{E}_0\left[\int_0^T \text{IM}(\partial_{Q(t)} V(t))\, dt\right]$$

- IM is additional collateral to mitigate counterparty risk over MPoR ($\sim$ 10D)
- Bilateral IM: both c/parties post to 3^{rd}-party custodians $\implies$ needs funding
- In practice, portfolio hedges attract bilateral &/or clearing-house IM too
- MVA reflects funding costs in valuations $\implies$ spectre of FVA debate

[2] See Green and Kenyon ('15) for detailed derivation

Motivation for IM

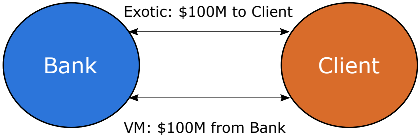

Figure: Exposure, variation margin and initial margin

Motivation for IM

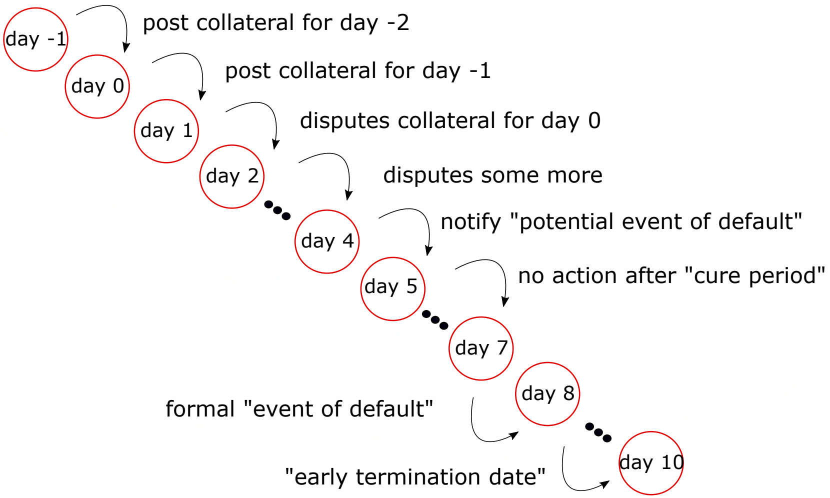

Figure: Event sequence during the margin period of risk: *a la* Andersen *et al.* ('17)

MVA: Motivation and Logistics 1

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR[2]

$$\text{MVA} = \mathbb{E}_0\left[\int_0^T \text{IM}(\partial_{Q(t)} V(t))\, dt\right]$$

- IM is additional collateral to mitigate counterparty risk over MPoR ($\sim$ 10D)
- Bilateral IM: both c/parties post to 3^{rd}-party custodians $\implies$ needs funding
- In practice, portfolio hedges attract bilateral &/or clearing-house IM too
- MVA reflects funding costs in valuations $\implies$ spectre of FVA debate

[2] See Green and Kenyon ('15) for detailed derivation

Motivation for IM

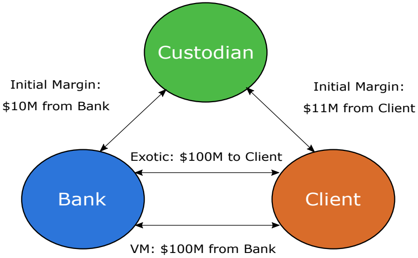

Figure: Exposure, variation margin and initial margin

Funding IM

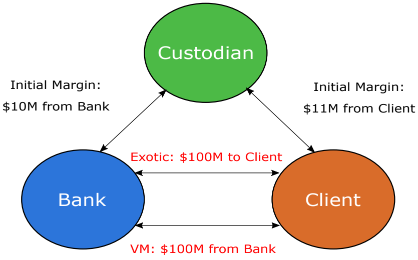

Figure: Exposure, variation margin and initial margin

Funding IM

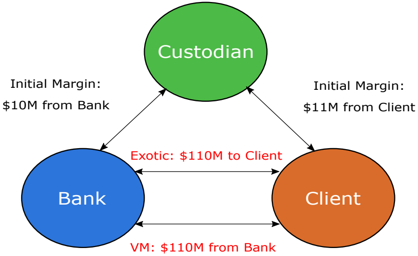

Figure: Exposure, variation margin and initial margin

Funding IM

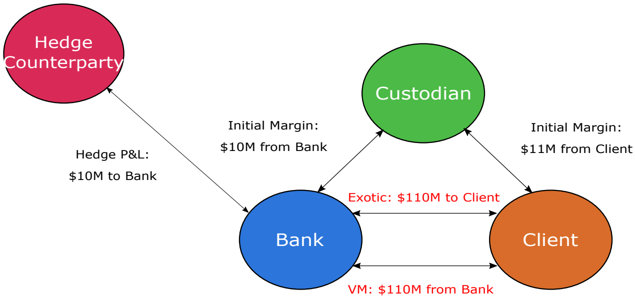

Figure: Exposure, variation margin and initial margin

Funding IM

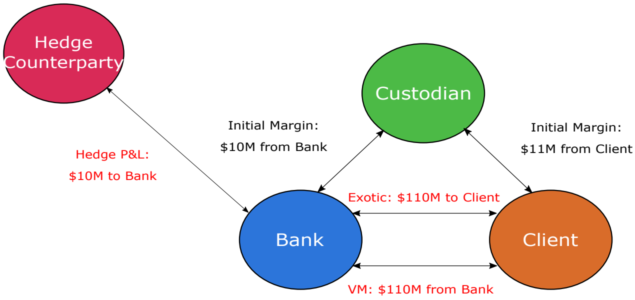

Figure: Exposure, variation margin and initial margin

MVA: Motivation and Logistics 1

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR[2]

$$\text{MVA} = \mathbb{E}_0\left[\int_0^T \text{IM}(\partial_{Q(t)} V(t))\, dt\right]$$

- IM is additional collateral to mitigate counterparty risk over MPoR ($\sim$ 10D)
- Bilateral IM: both c/parties post to 3^{rd}-party custodians $\implies$ needs funding
- In practice, portfolio hedges attract bilateral &/or clearing-house IM too
- MVA reflects funding costs in valuations $\implies$ spectre of FVA debate

[2] See Green and Kenyon ('15) for detailed derivation

Full Trade Impact on IM Requirements

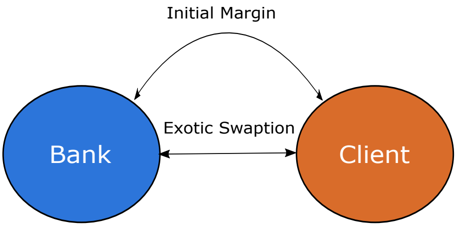

Figure: IM due to client trade and hedge trade/s

Full Trade Impact on IM Requirements

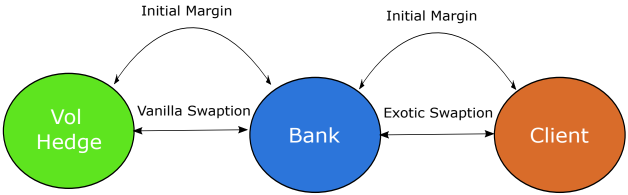

Figure: IM due to client trade and hedge trade/s

Full Trade Impact on IM Requirements

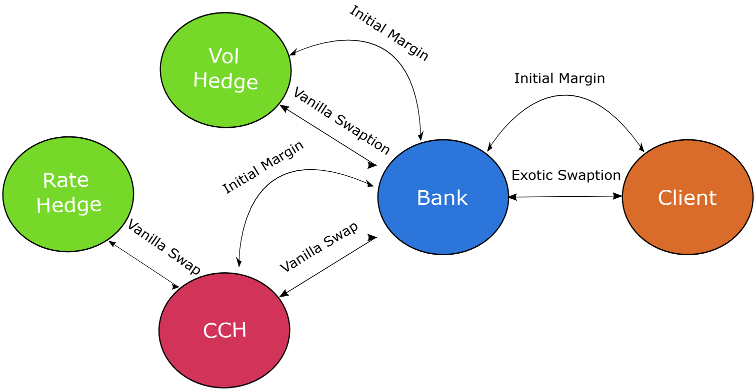

Figure: IM due to client trade and hedge trade/s

MVA: Motivation and Logistics 1

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR[2]

$$\text{MVA} = \mathbb{E}_0\left[\int_0^T \text{IM}(\partial_{Q(t)} V(t))\, dt\right]$$

- IM is additional collateral to mitigate counterparty risk over MPoR ($\sim$ 10D)
- Bilateral IM: both c/parties post to 3^{rd}-party custodians $\Longrightarrow$ needs funding
- In practice, portfolio hedges attract bilateral &/or clearing-house IM too
- MVA reflects funding costs in valuations $\Longrightarrow$ spectre of FVA debate

[2] See Green and Kenyon ('15) for detailed derivation

MVA: Motivation and Logistics 2

- ISDA proposed a sensitivity-based approximation to 99%-10D VaR for IM
- Sensitivities over *eg*. swap rates & implied vols, $Q = [S_1, \ldots, S_{N_S}, \nu_1, \ldots, \nu_{N_\nu}]$

$$\mathrm{IM}_{\mathrm{Delta}} \approx \sqrt{\partial_S' V \, \Sigma \, \partial_S V}$$

- Typical to use Jacobians to obtain Q-sensitivities from θ-sensitivities
- This just translates risk over $f_1, \sigma_1, \ldots$ to risk over $S_1, \nu_1, \ldots$

$$\partial_Q V = \partial_\theta V \, (\partial_\theta Q)^{-1}$$

- What if $N_\theta \neq N_Q$? $N_\theta < N_Q \rightarrow$ pseudo-inverse, $N_\theta > N_Q \rightarrow$ bucketing[2]
- The $N_\theta > N_Q$ case will enforced by model design and bucketing will be used[3]

---

[2] Could also use sophisticated shape-weighted bucketing, risk curves, *etc*.
[3] Fries ('18) may have an alternative for this

MVA: Motivation and Logistics 2

- ISDA proposed a sensitivity-based approximation to 99%-10D VaR for IM
- Sensitivities over *eg*. swap rates & implied vols, $Q = [S_1, \ldots, S_{N_S}, \nu_1, \ldots, \nu_{N_\nu}]$

$$\mathrm{IM}_{\mathrm{Delta}} \approx \sqrt{\partial_S' V \, \Sigma \, \partial_S V}$$

- Typical to use Jacobians to obtain Q-sensitivities from θ-sensitivities
- This just translates risk over $f_1, \sigma_1, \ldots$ to risk over $S_1, \nu_1, \ldots$

$$\partial_Q V = \partial_\theta V \, (\partial_\theta Q)^{-1}$$

- What if $N_\theta \neq N_Q$? $N_\theta < N_Q \rightarrow$ pseudo-inverse, $N_\theta > N_Q \rightarrow$ bucketing[2]
- The $N_\theta > N_Q$ case will enforced by model design and bucketing will be used[3]

[2] Could also use sophisticated shape-weighted bucketing, risk curves, *etc*.
[3] Fries ('18) may have an alternative for this

MVA: Motivation and Logistics 2

- ISDA proposed a sensitivity-based approximation to 99%-10D VaR for IM
- Sensitivities over *eg*. swap rates & implied vols, $Q = [S_1, \ldots, S_{N_S}, \nu_1, \ldots, \nu_{N_\nu}]$

$$\mathrm{IM}_{\mathrm{Vega}} \approx \sqrt{\partial_\nu' V \, \Omega \, \partial_\nu V}$$

- Typical to use Jacobians to obtain Q-sensitivities from θ-sensitivities
- This just translates risk over $f_1, \sigma_1, \ldots$ to risk over $S_1, \nu_1, \ldots$

$$\partial_Q V = \partial_\theta V \left(\partial_\theta Q\right)^{-1}$$

- What if $N_\theta \neq N_Q$? $N_\theta < N_Q \;\rightarrow$ pseudo-inverse, $N_\theta > N_Q \;\rightarrow$ bucketing[2]
- The $N_\theta > N_Q$ case will enforced by model design and bucketing will be used[3]

[2] Could also use sophisticated shape-weighted bucketing, risk curves, *etc*.
[3] Fries ('18) may have an alternative for this

MVA: Motivation and Logistics 2

- ISDA proposed a sensitivity-based approximation to 99%-10D VaR for IM
- Sensitivities over *eg*. swap rates & implied vols, $Q = [S_1, \ldots, S_{N_S}, \nu_1, \ldots, \nu_{N_\nu}]$

$$\text{IM}_{\text{Delta}} \approx \sqrt{\partial_S' V \, \Sigma \, \partial_S V}$$

- Typical to use Jacobians to obtain Q-sensitivities from θ-sensitivities
- This just translates risk over $f_1, \sigma_1, \ldots$ to risk over $S_1, \nu_1, \ldots$

$$\partial_Q V = \partial_\theta V \, (\partial_\theta Q)^{-1}$$

- What if $N_\theta \neq N_Q$? $N_\theta < N_Q \;\rightarrow$ pseudo-inverse, $N_\theta > N_Q \;\rightarrow$ bucketing[2]
- The $N_\theta > N_Q$ case will enforced by model design and bucketing will be used[3]

[2] Could also use sophisticated shape-weighted bucketing, risk curves, *etc*.
[3] Fries ('18) may have an alternative for this

MVA: Motivation and Logistics 2

- ISDA proposed a sensitivity-based approximation to 99%-10D VaR for IM
- Sensitivities over *eg*. swap rates & implied vols, $Q = [S_1, \ldots, S_{N_S}, \nu_1, \ldots, \nu_{N_\nu}]$

$$\text{IM}_{\text{Delta}} \approx \sqrt{\partial_S' V \, \Sigma \, \partial_S V}$$

- Typical to use Jacobians to obtain Q-sensitivities from θ-sensitivities
- This just translates risk over $f_1, \sigma_1, \ldots$ to risk over $S_1, \nu_1, \ldots$

$$\partial_Q V = \partial_\theta V \, (\partial_\theta Q)^{-1}$$

- What if $N_\theta \neq N_Q$? $N_\theta < N_Q \rightarrow$ pseudo-inverse, $N_\theta > N_Q \rightarrow$ bucketing[2]
- The $N_\theta > N_Q$ case will enforced by model design and bucketing will be used[3]

[2] Could also use sophisticated shape-weighted bucketing, risk curves, *etc*.
[3] Fries ('18) may have an alternative for this

MVA: Motivation and Logistics 2

- ISDA proposed a sensitivity-based approximation to 99%-10D VaR for IM
- Sensitivities over *eg*. swap rates & implied vols, $Q = [S_1, \ldots, S_{N_S}, \nu_1, \ldots, \nu_{N_\nu}]$

$$\mathrm{IM}_{\mathrm{Delta}} \approx \sqrt{\partial_S' V \, \Sigma \, \partial_S V}$$

- Typical to use Jacobians to obtain Q-sensitivities from θ-sensitivities
- This just translates risk over $f_1, \sigma_1, \ldots$ to risk over $S_1, \nu_1, \ldots$

$$\partial_Q V = \partial_\theta V \left(\partial_\theta Q\right)^{-1}$$

- What if $N_\theta \neq N_Q$? $N_\theta < N_Q \rightarrow$ pseudo-inverse, $N_\theta > N_Q \rightarrow$ bucketing[2]
- The $N_\theta > N_Q$ case will enforced by model design and bucketing will be used[3]

[2] Could also use sophisticated shape-weighted bucketing, risk curves, *etc*.
[3] Fries ('18) may have an alternative for this

Bucketing

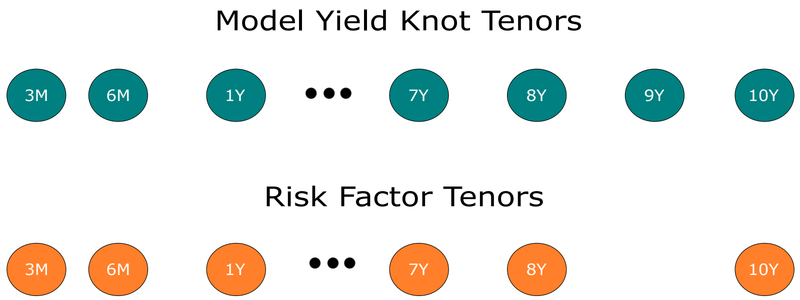

Figure: Bucketing to ensure invertible Jacocbians

Bucketing

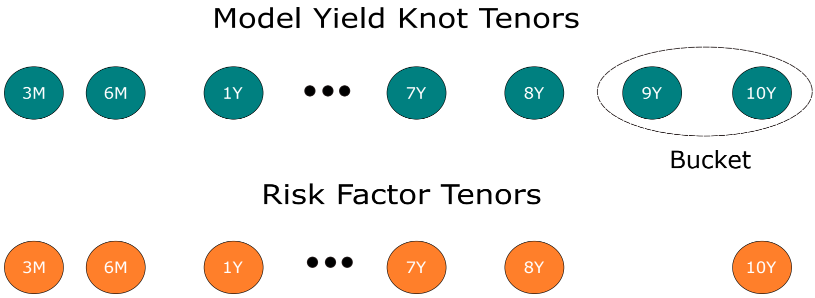

Figure: Bucketing to ensure invertible Jacocbians

MVA: Motivation and Logistics 2

- ISDA proposed a sensitivity-based approximation to 99%-10D VaR for IM
- Sensitivities over *eg*. swap rates & implied vols, $Q = [S_1, \ldots, S_{N_S}, \nu_1, \ldots, \nu_{N_\nu}]$

$$\mathrm{IM}_{\mathrm{Delta}} \approx \sqrt{\partial_S' V \, \Sigma \, \partial_S V}$$

- Typical to use Jacobians to obtain Q-sensitivities from θ-sensitivities
- This just translates risk over $f_1, \sigma_1, \ldots$ to risk over $S_1, \nu_1, \ldots$

$$\partial_{Q(t)} V(t) = \partial_\theta V(t) \, (\partial_\theta Q(t))^{-1}$$

- What if $N_\theta \neq N_Q$? $N_\theta < N_Q \; \rightarrow$ pseudo-inverse, $N_\theta > N_Q \; \rightarrow$ bucketing[2]
- The $N_\theta > N_Q$ case will enforced by model design and bucketing will be used[3]

---

[2] Could also use sophisticated shape-weighted bucketing, risk curves, *etc*.
[3] Fries ('18) may have an alternative for this

Bucketing

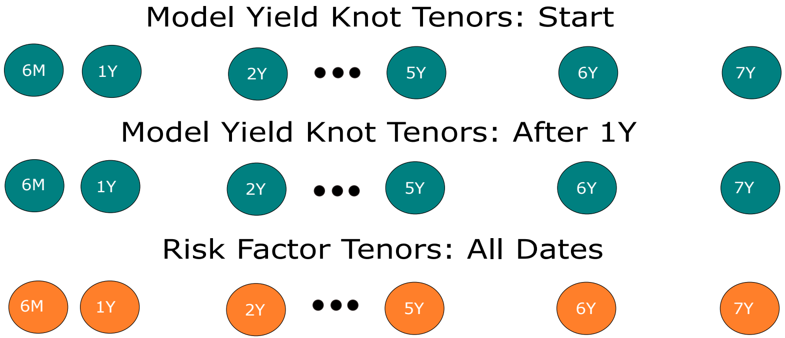

Figure: Bucketing to ensure invertible future Jacocbians

Bucketing

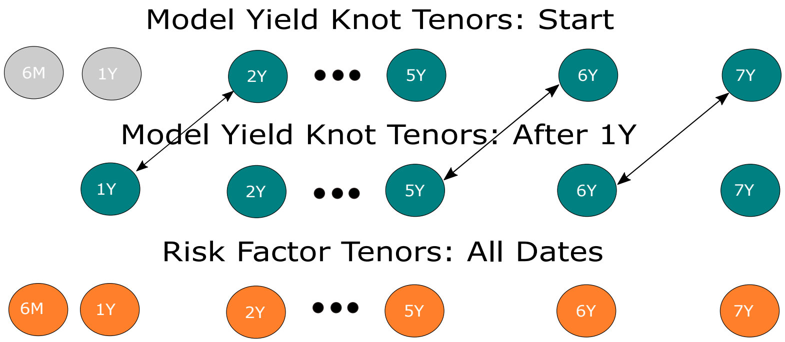

Figure: Bucketing to ensure invertible future Jacocbians

Bucketing

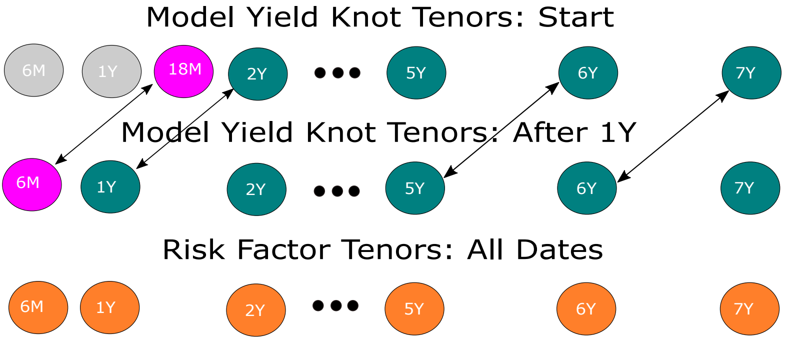

Figure: Bucketing to ensure invertible future Jacocbians

Bucketing

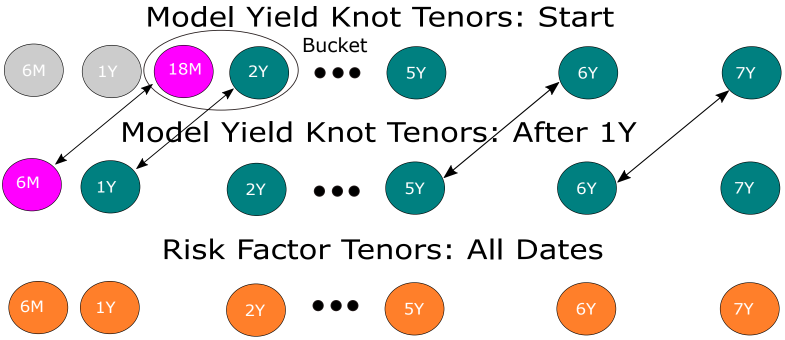

Figure: Bucketing to ensure invertible future Jacocbians

Presentation Outline

- CVA Greeks and MVA via “Future” Greeks
- Future Greeks as a by-product of AD-on-LSMC
- AD efficiencies for LSMC: large-sample regression coefficient dependencies

Presentation Outline

- CVA Greeks and MVA via "Future" Greeks
- Future Greeks as a by-product of AD-on-LSMC
- AD efficiencies for LSMC: large-sample regression coefficient dependencies

Accuracy of Future Greeks from LSMC 1

- Our $\hat{V}_i$ come from regressing $\hat{V}_{i+1}$ onto N_B basis functions $\phi(X_i)$

$$\hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- Can establish MSE of LSMC error in $\hat{V}_{p,i}$

$$\begin{aligned} \text{MSE}(\hat{V}_{p,i}|X_i) &= \mathbb{E}[(\hat{V}_{p,i} - V_{p,i})^2|X_i] \\ &= \phi(X_{p,i})'\text{var}(\hat{\beta}|X_i)\phi(X_{p,i}) + (V_{p,i} - \phi(X_{p,i}) \cdot \beta_\infty)^2 \end{aligned}$$

- Is the basis good for $\partial_\theta \hat{V}_{i+1}$? How does the bias react? Need more flexibility?
- What about the variance of $\partial_\theta \hat{V}_{i+1}$? Need larger N_P?

Accuracy of Future Greeks from LSMC 1

- Our $\hat{V}_i$ come from regressing $\hat{V}_{i+1}$ onto N_B basis functions $\phi(X_i)$

$$\hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$

$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- Can establish MSE of LSMC error in $\hat{V}_{p,i}$

$$\mathrm{MSE}(\hat{V}_{p,i}|X_i) = \mathbb{E}[(\hat{V}_{p,i} - V_{p,i})^2|X_i]$$

$$= \phi(X_{p,i})'\mathrm{var}(\hat{\beta}|X_i)\phi(X_{p,i}) + (V_{p,i} - \phi(X_{p,i}) \cdot \beta_\infty)^2$$

- Is the basis good for $\partial_\theta \hat{V}_{i+1}$? How does the bias react? Need more flexibility?
- What about the variance of $\partial_\theta \hat{V}_{i+1}$? Need larger N_P?

Accuracy of Future Greeks from LSMC 1

- Our $\hat{V}_i$ come from regressing $\hat{V}_{i+1}$ onto N_B basis functions $\phi(X_i)$

$$\hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$
$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- Can establish MSE of LSMC error in $\hat{V}_{p,i}$

$$\mathrm{MSE}(\hat{V}_{p,i}|X_i) = \mathbb{E}[((\hat{V}_{p,i} - \phi(X_{p,i}) \cdot \beta_\infty) - (V_{p,i} - \phi(X_{p,i}) \cdot \beta_\infty))^2|X_i]$$
$$= \phi(X_{p,i})'\mathrm{var}(\hat{\beta}|X_i)\phi(X_{p,i}) + (V_{p,i} - \phi(X_{p,i}) \cdot \beta_\infty)^2$$

- Is the basis good for $\partial_\theta \hat{V}_{i+1}$? How does the bias react? Need more flexibility?
- What about the variance of $\partial_\theta \hat{V}_{i+1}$? Need larger N_P?

Accuracy of Future Greeks from LSMC 1

- Our $\hat{V}_i$ come from regressing $\hat{V}_{i+1}$ onto N_B basis functions $\phi(X_i)$

$$\begin{aligned}\hat{V}_{p,i} &= \phi(X_{i,p}) \cdot \hat{\beta} \\ \hat{\beta} &= (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}\end{aligned}$$

- Can establish MSE of LSMC error in $\hat{V}_{p,i}$

$$\begin{aligned}\text{MSE}(\hat{V}_{p,i}|X_i) &= \mathbb{E}[((\hat{V}_{p,i} - \phi(X_{p,i}) \cdot \beta_\infty) - (V_{p,i} - \phi(X_{p,i}) \cdot \beta_\infty))^2|X_i] \\ &\approx \phi(X_{p,i})'\text{var}(\hat{\beta}|X_i)\phi(X_{p,i}) + (V_{p,i} - \phi(X_{p,i}) \cdot \beta_\infty)^2\end{aligned}$$

- Is the basis good for $\partial_\theta \hat{V}_{i+1}$? How does the bias react? Need more flexibility?
- What about the variance of $\partial_\theta \hat{V}_{i+1}$? Need larger N_P?

Accuracy of Future Greeks from LSMC 1

- Our $\hat{V}_i$ come from regressing $\hat{V}_{i+1}$ onto N_B basis functions $\phi(X_i)$

$$\hat{V}_{p,i} = \phi(X_{i,p}) \cdot \hat{\beta}$$
$$\hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}$$

- Can establish MSE of LSMC error in $\hat{V}_{p,i}$

$$\begin{aligned}\text{MSE}(\hat{V}_{p,i}|X_i) &= \mathbb{E}[((\hat{V}_{p,i} - \phi(X_{p,i}) \cdot \beta_\infty) - (V_{p,i} - \phi(X_{p,i}) \cdot \beta_\infty))^2 |X_i] \\ &= \phi(X_{p,i})'\text{var}(\hat{\beta}|X_i)\phi(X_{p,i}) + (V_{p,i} - \phi(X_{p,i}) \cdot \beta_\infty)^2\end{aligned}$$

- Is the basis good for $\partial_\theta \hat{V}_{i+1}$? How does the bias react? Need more flexibility?
- What about the variance of $\partial_\theta \hat{V}_{i+1}$? Need larger N_P?

numerix

Accuracy of Future Greeks from LSMC 1

- Our $\hat{V}_i$ come from regressing $\hat{V}_{i+1}$ onto N_B basis functions $\phi(X_i)$

$$\begin{aligned} \partial_\theta \hat{V}_{p,i} &= \phi(X_{i,p}) \cdot \partial_\theta \hat{\beta} \\ \partial_\theta \hat{\beta} &= (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1} \end{aligned}$$

- Can establish MSE of LSMC error in $\hat{V}_{p,i}$

$$\begin{aligned} \mathrm{MSE}(\partial_\theta \hat{V}_{p,i}) &= \mathbb{E}[((\partial_\theta \hat{V}_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty) - (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty))^2] \\ &= \phi(X_{p,i})'\mathrm{var}(\partial_\theta \hat{\beta})\phi(X_{p,i}) + (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty)^2 \end{aligned}$$

- Is the basis good for $\partial_\theta \hat{V}_{i+1}$? How does the bias react? Need more flexibility?
- What about the variance of $\partial_\theta \hat{V}_{i+1}$? Need larger N_P?

Accuracy of Future Greeks from LSMC 1

- Our $\hat{V}_i$ come from regressing $\hat{V}_{i+1}$ onto N_B basis functions $\phi(X_i)$

$$\begin{aligned} \partial_\theta \hat{V}_{p,i} &= \phi(X_{i,p}) \cdot \partial_\theta \hat{\beta} \\ \partial_\theta \hat{\beta} &= (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1} \end{aligned}$$

- Can establish MSE of LSMC error in $\hat{V}_{p,i}$

$$\begin{aligned} \mathrm{MSE}(\partial_\theta \hat{V}_{p,i}) &= \mathbb{E}[((\partial_\theta \hat{V}_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty) - (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty))^2] \\ &= \phi(X_{p,i})'\mathrm{var}(\,\partial_\theta \hat{\beta}\,)\phi(X_{p,i}) + (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty)^2 \end{aligned}$$

- Is the basis good for $\partial_\theta \hat{V}_{i+1}$? How does the bias react? Need more flexibility?
- What about the variance of $\partial_\theta \hat{V}_{i+1}$? Need larger N_P?

Accuracy of Future Greeks from LSMC 1

- Our $\hat{V}_i$ come from regressing $\hat{V}_{i+1}$ onto N_B basis functions $\phi(X_i)$

$$\begin{aligned}\partial_\theta \hat{V}_{p,i} &= \phi(X_{i,p}) \cdot \partial_\theta \hat{\beta} \\ \partial_\theta \hat{\beta} &= (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}\end{aligned}$$

- Can establish MSE of LSMC error in $\hat{V}_{p,i}$

$$\begin{aligned}\mathrm{MSE}(\partial_\theta \hat{V}_{p,i}) &= \mathbb{E}[((\partial_\theta \hat{V}_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty) - (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty))^2] \\ &= \phi(X_{p,i})'\mathrm{var}(\,\partial_\theta \hat{\beta}\,)\phi(X_{p,i}) + (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty)^2\end{aligned}$$

- Is the basis good for $\partial_\theta \hat{V}_{i+1}$? How does the bias react? Need more flexibility?
- What about the variance of $\partial_\theta \hat{V}_{i+1}$? Need larger N_P?

Accuracy of Future Greeks from LSMC 1

- Our $\hat{V}_i$ come from regressing $\hat{V}_{i+1}$ onto N_B basis functions $\phi(X_i)$

$$\partial_\theta \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \partial_\theta \hat{\beta}$$

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can establish MSE of LSMC error in $\hat{V}_{p,i}$

$$\begin{aligned} \mathrm{MSE}(\partial_\theta \hat{V}_{p,i}) &= \mathbb{E}[((\partial_\theta \hat{V}_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty) - (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty))^2] \\ &= \phi(X_{p,i})' \mathrm{var}(\partial_\theta \hat{\beta}) \phi(X_{p,i}) + (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty)^2 \end{aligned}$$

- Is the basis good for $\partial_\theta \hat{V}_{i+1}$? How does the bias react? Need more flexibility?
- What about the variance of $\partial_\theta \hat{V}_{i+1}$? Need larger N_P?

Accuracy of Future Greeks from LSMC 1

- Our $\hat{V}_i$ come from regressing $\hat{V}_{i+1}$ onto N_B basis functions $\phi(X_i)$

$$\partial_\theta \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \partial_\theta \hat{\beta}$$

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Can establish MSE of LSMC error in $\hat{V}_{p,i}$

$$\begin{aligned}\mathrm{MSE}(\partial_\theta \hat{V}_{p,i}) &= \mathbb{E}[((\partial_\theta \hat{V}_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty) - (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty))^2] \\ &= \phi(X_{p,i})'\mathrm{var}(\partial_\theta \hat{\beta})\phi(X_{p,i}) + (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty)^2\end{aligned}$$

- Is the basis good for $\partial_\theta \hat{V}_{i+1}$? How does the bias react? Need more flexibility?
- What about the variance of $\partial_\theta \hat{V}_{i+1}$? Need larger N_P?

Accuracy of Future Greeks from LSMC 1

- Our $\hat{V}_i$ come from regressing $\hat{V}_{i+1}$ onto N_B basis functions $\phi(X_i)$

$$\partial_\theta \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \partial_\theta \hat{\beta}$$

$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\mathrm{var}(\partial_\theta \hat{V}_{i+1})\phi(X_i)(\phi(X_i)'\phi(X_i))^{-1}$$

- Can establish MSE of LSMC error in $\hat{V}_{p,i}$

$$\begin{aligned}\mathrm{MSE}(\partial_\theta \hat{V}_{p,i}) &= \mathbb{E}[((\partial_\theta \hat{V}_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty) - (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty))^2] \\ &= \phi(X_{p,i})'\mathrm{var}(\partial_\theta \hat{\beta})\phi(X_{p,i}) + (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty)^2\end{aligned}$$

- Is the basis good for $\partial_\theta \hat{V}_{i+1}$? How does the bias react? Need more flexibility?
- What about the variance of $\partial_\theta \hat{V}_{i+1}$? Need larger N_P?

Accuracy of Future Greeks from LSMC 1

- Our $\hat{V}_i$ come from regressing $\hat{V}_{i+1}$ onto N_B basis functions $\phi(X_i)$

$$\begin{aligned} \partial_\theta \hat{V}_{p,i} &= \phi(X_{i,p}) \cdot \partial_\theta \hat{\beta} \\ \partial_\theta \hat{\beta} &= (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\mathrm{var}(\partial_\theta \hat{V}_{i+1})\phi(X_i)(\phi(X_i)'\phi(X_i))^{-1} \end{aligned}$$

- Can establish MSE of LSMC error in $\hat{V}_{p,i}$

$$\begin{aligned} \mathrm{MSE}(\partial_\theta \hat{V}_{p,i}) &= \mathbb{E}[((\partial_\theta \hat{V}_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty) - (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty))^2] \\ &= \phi(X_{p,i})'\mathrm{var}(\partial_\theta \hat{\beta})\phi(X_{p,i}) + (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty)^2 \end{aligned}$$

- Is the basis good for $\partial_\theta \hat{V}_{i+1}$? How does the bias react? Need more flexibility?
- What about the variance of $\partial_\theta \hat{V}_{i+1}$? Need larger N_P?

Accuracy of Future Greeks from LSMC 1

- Our $\hat{V}_i$ come from regressing $\hat{V}_{i+1}$ onto N_B basis functions $\phi(X_i)$

$$\partial_\theta \hat{V}_{p,i} = \phi(X_{i,p}) \cdot \partial_\theta \hat{\beta}$$
$$\partial_\theta \hat{\beta} = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\text{var}(\partial_\sigma \hat{V}_{i+1})\phi(X_i)(\phi(X_i)'\phi(X_i))^{-1}$$

- Can establish MSE of LSMC error in $\hat{V}_{p,i}$

$$\text{MSE}(\partial_\theta \hat{V}_{p,i}) = \mathbb{E}[((\partial_\theta \hat{V}_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty) - (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty))^2]$$
$$= \phi(X_{p,i})'\text{var}(\partial_\theta \hat{\beta})\phi(X_{p,i}) + (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty)^2$$

- Is the basis good for $\partial_\theta \hat{V}_{i+1}$? How does the bias react? Need more flexibility?
- What about the variance of $\partial_\theta \hat{V}_{i+1}$? Need larger N_P?

Accuracy of Future Greeks from LSMC 1

- Our $\hat{V}_i$ come from regressing $\hat{V}_{i+1}$ onto N_B basis functions $\phi(X_i)$

$$\begin{aligned}\partial_\theta \hat{V}_{p,i} &= \phi(X_{i,p}) \cdot \partial_\theta \hat{\beta} \\ \partial_\theta \hat{\beta} &= (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\text{var}(\cdots dW(t))\phi(X_i)(\phi(X_i)'\phi(X_i))^{-1}\end{aligned}$$

- Can establish MSE of LSMC error in $\hat{V}_{p,i}$

$$\begin{aligned}\text{MSE}(\partial_\theta \hat{V}_{p,i}) &= \mathbb{E}[((\partial_\theta \hat{V}_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty) - (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty))^2] \\ &= \phi(X_{p,i})'\text{var}(\partial_\theta \hat{\beta})\phi(X_{p,i}) + (\partial_\theta V_{p,i} - \phi(X_{p,i}) \cdot \partial_\theta \beta_\infty)^2\end{aligned}$$

- Is the basis good for $\partial_\theta \hat{V}_{i+1}$? How does the bias react? Need more flexibility?
- What about the variance of $\partial_\theta \hat{V}_{i+1}$? Need larger N_P?

numerix

AD-on-LSMC Accuracy

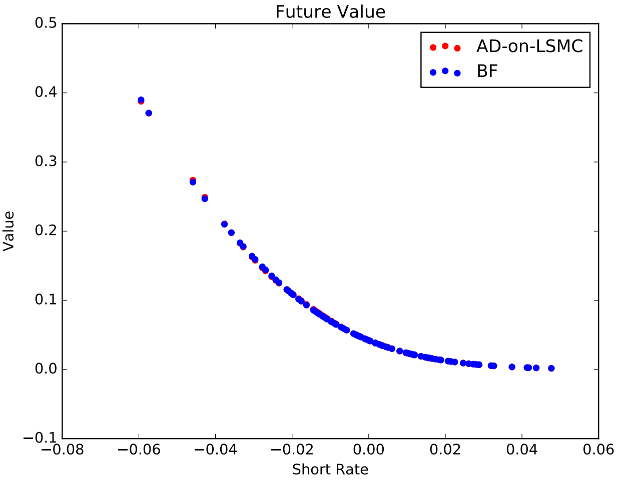

Figure: AD-on-LSMC Values *vs.* Brute-Force: 10-into-16 Bermudan at 5Y Observation

AD-on-LSMC Accuracy

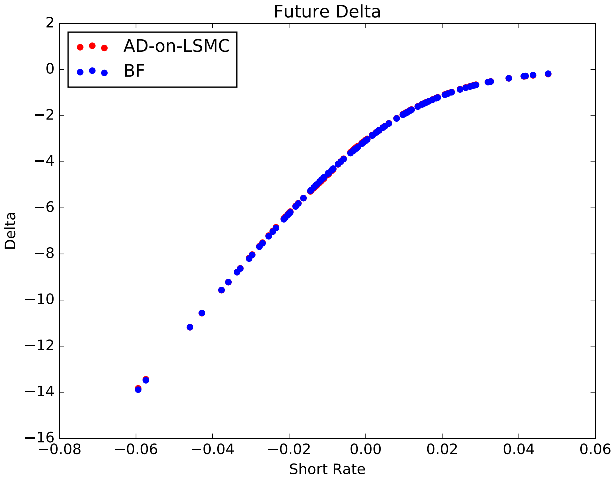

Figure: AD-on-LSMC Deltas *vs.* Brute-Force: 10-into-16 Bermudan at 5Y Observation

AD-on-LSMC Accuracy

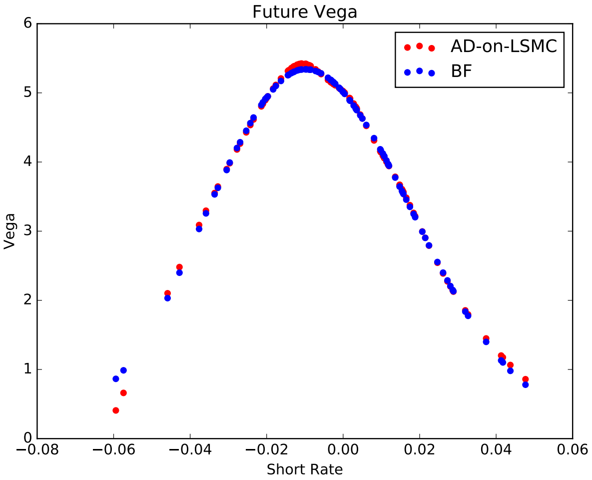

Figure: AD-on-LSMC Vegas *vs.* Brute-Force: 10-into-16 Bermudan at 5Y Observation

Accuracy of Future Greeks from LSMC 2

- *Many* engineering techniques available to improve LSMC accuracy

1. Craft basis on a trade-by-trade basis and incorporate functions of θ

$$V(X_{p,i},\theta) \approx \beta_0 + \beta_1 V^{euro}(X_{p,i},\theta) + \beta_2 V^{euro}(X_{p,i},\theta)\, w(X_{p,i},\theta) + \cdots$$

2. Use control variates to reduce variance in V_{i+1}

$$\hat{V}_{p,i+1} = \phi(X_{p,i}) \cdot \beta + \epsilon_{p,i}$$

3. Assess impact of using $\hat{V}_{i+1}$ *vs.* C_{i+1,N_T} as regressands: bias *vs.* variance

- As for LSMC exposures, need engineering & validation in complex cases

Accuracy of Future Greeks from LSMC 2

- *Many* engineering techniques available to improve LSMC accuracy

1. Craft basis on a trade-by-trade basis and incorporate functions of θ

$$V(X_{p,i},\theta) \approx \beta_0 + \beta_1\, V^{euro}(X_{p,i},\theta) + \beta_2\, V^{euro}(X_{p,i},\theta)\, w(X_{p,i},\theta) + \cdots$$

2. Use control variates to reduce variance in V_{i+1}

$$\hat{V}_{p,i+1} = \phi(X_{p,i}) \cdot \beta + \epsilon_{p,i}$$

3. Assess impact of using $\hat{V}_{i+1}$ *vs.* C_{i+1,N_T} as regressands: bias *vs.* variance

- As for LSMC exposures, need engineering & validation in complex cases

Accuracy of Future Greeks from LSMC 2

- *Many* engineering techniques available to improve LSMC accuracy

1. Craft basis on a trade-by-trade basis and incorporate functions of θ

$$V(X_{p,i},\theta) \approx \beta_0 + \beta_1\, V^{euro}(X_{p,i},\theta) + \beta_2\, V^{euro}(X_{p,i},\theta)\, w(X_{p,i},\theta) + \cdots$$

2. Use control variates to reduce variance in V_{i+1}

$$\hat{V}_{p,i+1} = \phi(X_{p,i}) \cdot \beta + \epsilon_{p,i}$$

3. Assess impact of using $\hat{V}_{i+1}$ *vs.* C_{i+1,N_T} as regressands: bias *vs.* variance

- As for LSMC exposures, need engineering & validation in complex cases

Accuracy of Future Greeks from LSMC 2

- *Many* engineering techniques available to improve LSMC accuracy

1. Craft basis on a trade-by-trade basis and incorporate functions of θ

$$V(X_{p,i},\theta) \approx \beta_0 + \beta_1 V^{euro}(X_{p,i},\theta) + \beta_2 V^{euro}(X_{p,i},\theta)\, w(X_{p,i},\theta) + \cdots$$

2. Use control variates to reduce variance in V_{i+1}

$$\hat{V}_{p,i+1} - \upsilon_{p,i} = \phi(X_{p,i}) \cdot \beta + \epsilon_{p,i}$$

3. Assess impact of using $\hat{V}_{i+1}$ *vs.* C_{i+1,N_T} as regressands: bias *vs.* variance

- As for LSMC exposures, need engineering & validation in complex cases

Accuracy of Future Greeks from LSMC 2

- *Many* engineering techniques available to improve LSMC accuracy

1. Craft basis on a trade-by-trade basis and incorporate functions of θ

$$V(X_{p,i},\theta) \approx \beta_0 + \beta_1\, V^{euro}(X_{p,i},\theta) + \beta_2\, V^{euro}(X_{p,i},\theta)\, w(X_{p,i},\theta) + \cdots$$

2. Use control variates to reduce variance in V_{i+1}

$$\hat{V}_{p,i+1} - \upsilon_{p,i} = \phi(X_{p,i}) \cdot \beta + \epsilon^{\upsilon}_{p,i}$$

3. Assess impact of using $\hat{V}_{i+1}$ *vs.* C_{i+1,N_T} as regressands: bias *vs.* variance

- As for LSMC exposures, need engineering & validation in complex cases

Accuracy of Future Greeks from LSMC 2

- *Many* engineering techniques available to improve LSMC accuracy

1. Craft basis on a trade-by-trade basis and incorporate functions of θ

$$V(X_{p,i},\theta) \approx \beta_0 + \beta_1\, V^{euro}(X_{p,i},\theta) + \beta_2\, V^{euro}(X_{p,i},\theta)\, w(X_{p,i},\theta) + \cdots$$

2. Use control variates to reduce variance in V_{i+1}

$$\hat{V}_{p,i+1} - V^{euro}_{p,i+1} - V^{euro}_{p,i} = \phi(X_{p,i}) \cdot \beta + \epsilon^{\upsilon}_{p,i}$$

3. Assess impact of using $\hat{V}_{i+1}$ *vs.* C_{i+1,N_T} as regressands: bias *vs.* variance

- As for LSMC exposures, need engineering & validation in complex cases

Accuracy of Future Greeks from LSMC 2

- *Many* engineering techniques available to improve LSMC accuracy

1. Craft basis on a trade-by-trade basis and incorporate functions of θ

$$V(X_{p,i},\theta) \approx \beta_0 + \beta_1 V^{euro}(X_{p,i},\theta) + \beta_2 V^{euro}(X_{p,i},\theta)\, w(X_{p,i},\theta) + \cdots$$

2. Use control variates to reduce variance in V_{i+1}

$$\hat{V}_{p,i+1} = \phi(X_{p,i}) \cdot \beta + \epsilon_{p,i}$$

3. Assess impact of using $\hat{V}_{i+1}$ *vs.* C_{i+1,N_T} as regressands: bias *vs.* variance

- As for LSMC exposures, need engineering & validation in complex cases

Alternative to AD-on-LSMC: Direct Greek Regression

- High-dimensional models, path-dependent products, complex payoffs *etc*.
- Can expect performance of LSMC Greeks to suffer, need alternative
- Can regress $\partial_{\theta_n} C_{i+1,N_T}$ directly onto dedicated basis, $\phi_{\theta_n}(X_i, \theta)$

$$\partial_{\theta_n} V_{p,i} = \mathbb{E}\left[\partial_{\theta_n} C_{i+1} | X_{p,i}\right] \longrightarrow \partial_{\theta_n} \hat{V}_{p,i} = \phi_{\theta_n}(X_i, \theta) \cdot \hat{\gamma}_{\theta_n}$$

- Main benefit is that basis only has to tailor to $\partial_{\theta_n} V_i$, not V_i & $\partial_\theta V_i$
- Expensive: $\hat{\beta}$ differentiated N_θ times is cheaper than $\hat{\gamma}_{\theta_n}$ computing N_θ times
- Can mix-&-match, using AD-on-LSMC for all but difficult members of θ

Alternative to AD-on-LSMC: Direct Greek Regression

- High-dimensional models, path-dependent products, complex payoffs *etc*.
- Can expect performance of LSMC Greeks to suffer, need alternative
- Can regress $\partial_{\theta_n} C_{i+1,N_T}$ directly onto dedicated basis, $\phi_{\theta_n}(X_i, \theta)$

$$\partial_{\theta_n} V_{p,i} = \mathbb{E}\left[\partial_{\theta_n} C_{i+1} | X_{p,i}\right] \longrightarrow \partial_{\theta_n} \hat{V}_{p,i} = \phi_{\theta_n}(X_i, \theta) \cdot \hat{\gamma}_{\theta_n}$$

- Main benefit is that basis only has to tailor to $\partial_{\theta_n} V_i$, not V_i & $\partial_\theta V_i$
- Expensive: $\hat{\beta}$ differentiated N_θ times is cheaper than $\hat{\gamma}_{\theta_n}$ computing N_θ times
- Can mix-&-match, using AD-on-LSMC for all but difficult members of θ

Alternative to AD-on-LSMC: Direct Greek Regression

- High-dimensional models, path-dependent products, complex payoffs *etc*.
- Can expect performance of LSMC Greeks to suffer, need alternative
- Can regress $\partial_{\theta_n} C_{i+1,N_T}$ directly onto dedicated basis, $\phi_{\theta_n}(X_i, \theta)$

$$\partial_{\theta_n} V_{p,i} = \mathbb{E}\left[\partial_{\theta_n} C_{i+1} | X_{p,i}\right] \longrightarrow \partial_{\theta_n} \hat{V}_{p,i} = \phi_{\theta_n}(X_i, \theta) \cdot \hat{\gamma}_{\theta_n}$$

- Main benefit is that basis only has to tailor to $\partial_{\theta_n} V_i$, not V_i & $\partial_\theta V_i$
- Expensive: $\hat{\beta}$ differentiated N_θ times is cheaper than $\hat{\gamma}_{\theta_n}$ computing N_θ times
- Can mix-&-match, using AD-on-LSMC for all but difficult members of θ

Alternative to AD-on-LSMC: Direct Greek Regression

- High-dimensional models, path-dependent products, complex payoffs *etc*.
- Can expect performance of LSMC Greeks to suffer, need alternative
- Can regress $\partial_{\theta_n} C_{i+1,N_T}$ directly onto dedicated basis, $\phi_{\theta_n}(X_i, \theta)$

$$\partial_{\theta_n} V_{p,i} = \mathbb{E}\left[\partial_{\theta_n} C_{i+1} | X_{p,i}\right] \longrightarrow \partial_{\theta_n} \hat{V}_{p,i} = \phi_{\theta_n}(X_i, \theta) \cdot \hat{\gamma}_{\theta_n}$$

- Main benefit is that basis only has to tailor to $\partial_{\theta_n} V_i$, not V_i & $\partial_\theta V_i$
- Expensive: $\hat{\beta}$ differentiated N_θ times is cheaper than $\hat{\gamma}_{\theta_n}$ computing N_θ times
- Can mix-&-match, using AD-on-LSMC for all but difficult members of θ

Alternative to AD-on-LSMC: Direct Greek Regression

- High-dimensional models, path-dependent products, complex payoffs *etc*.
- Can expect performance of LSMC Greeks to suffer, need alternative
- Can regress $\partial_{\theta_n} C_{i+1,N_T}$ directly onto dedicated basis, $\phi_{\theta_n}(X_i, \theta)$

$$\partial_{\theta_n} V_{p,i} = \mathbb{E}\left[\partial_{\theta_n} C_{i+1} | X_{p,i}\right] \longrightarrow \partial_{\theta_n} \hat{V}_{p,i} = \phi_{\theta_n}(X_i, \theta) \cdot \hat{\gamma}_{\theta_n}$$

- Main benefit is that basis only has to tailor to $\partial_{\theta_n} V_i$, not V_i & $\partial_\theta V_i$
- Expensive: $\hat{\beta}$ differentiated N_θ times is cheaper than $\hat{\gamma}_{\theta_n}$ computing N_θ times
- Can mix-&-match, using AD-on-LSMC for all but difficult members of θ

Alternative to AD-on-LSMC: Direct Greek Regression

- High-dimensional models, path-dependent products, complex payoffs *etc*.
- Can expect performance of LSMC Greeks to suffer, need alternative
- Can regress $\partial_{\theta_n} C_{i+1,N_T}$ directly onto dedicated basis, $\phi_{\theta_n}(X_i, \theta)$

$$\partial_{\theta_n} V_{p,i} = \mathbb{E}\left[\partial_{\theta_n} C_{i+1} | X_{p,i}\right] \longrightarrow \partial_{\theta_n} \hat{V}_{p,i} = \phi_{\theta_n}(X_i, \theta) \cdot \hat{\gamma}_{\theta_n}$$

- Main benefit is that basis only has to tailor to $\partial_{\theta_n} V_i$, not V_i & $\partial_\theta V_i$
- Expensive: $\hat{\beta}$ differentiated N_θ times is cheaper than $\hat{\gamma}_{\theta_n}$ computing N_θ times
- Can mix-&-match, using AD-on-LSMC for all but difficult members of θ

Presentation Outline

- CVA Greeks and MVA via "Future" Greeks
- Future Greeks as a by-product of AD-on-LSMC
- AD efficiencies for LSMC: large-sample regression coefficient dependencies

Presentation Outline

- CVA Greeks and MVA via "Future" Greeks
- Future Greeks as a by-product of AD-on-LSMC
- AD efficiencies for LSMC: large-sample regression coefficient dependencies

Coefficient Behavior and Dependencies in Large Samples

- Dependence upon θ gets propagated through the regression matrix

$$\partial_\theta \hat{\beta}_i = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Large-sample: ignore X_i-dependence in $\hat{\beta}$, & thus θ-dependence in X_i

$$\lim_{N_P\to\infty} \partial_{X_i}\hat{\beta}\,\partial_\theta X_i = \lim_{N_P\to\infty} \partial_{X_i}\left((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}\right)\partial_\theta X_i = 0$$

- Propagating through $\partial_{X_i}\hat{\beta}$ is as expensive as the main propagation of $\partial_\theta \hat{V}_{i+1}$
- Differentiating noise, $\partial_{X_i}\hat{\beta} = \partial_{X_i}(\beta_\infty - (\hat{\beta} - \beta_\infty)) = \partial_{X_i}(\hat{\beta} - \beta_\infty)) = \partial_{X_i}\epsilon_{\hat{\beta}}$
- Still important in presence of outliers/overfit, *eg*. in small samples

Coefficient Behavior and Dependencies in Large Samples

- Dependence upon θ gets propagated through the regression matrix

$$\partial_\theta \hat{\beta}_i = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Large-sample: ignore X_i-dependence in $\hat{\beta}$, & thus θ-dependence in X_i

$$\lim_{N_P\to\infty} \partial_{X_i}\hat{\beta}\,\partial_\theta X_i = \lim_{N_P\to\infty} \partial_{X_i}\left((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}\right)\partial_\theta X_i = 0$$

- Propagating through $\partial_{X_i}\hat{\beta}$ is as expensive as the main propagation of $\partial_\theta \hat{V}_{i+1}$
- Differentiating noise, $\partial_{X_i}\hat{\beta} = \partial_{X_i}(\beta_\infty - (\hat{\beta} - \beta_\infty)) = \partial_{X_i}(\hat{\beta} - \beta_\infty)) = \partial_{X_i}\epsilon_{\hat{\beta}}$
- Still important in presence of outliers/overfit, *eg*. in small samples

Coefficient Behavior and Dependencies in Large Samples

- Dependence upon θ gets propagated through the regression matrix

$$\partial_\theta \hat{\beta}_i = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Large-sample: ignore X_i-dependence in $\hat{\beta}$, & thus θ-dependence in X_i

$$\partial_{X_i}\hat{\beta}\,\partial_\theta X_i = \partial_{X_i}\left((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}\right)\partial_\theta X_i = 0$$

- Propagating through $\partial_{X_i}\hat{\beta}$ is as expensive as the main propagation of $\partial_\theta \hat{V}_{i+1}$
- Differentiating noise, $\partial_{X_i}\hat{\beta} = \partial_{X_i}(\beta_\infty - (\hat{\beta} - \beta_\infty)) = \partial_{X_i}(\hat{\beta} - \beta_\infty)) = \partial_{X_i}\epsilon_{\hat{\beta}}$
- Still important in presence of outliers/overfit, *eg*. in small samples

Coefficient Behavior and Dependencies in Large Samples

- Dependence upon θ gets propagated through the regression matrix

$$\partial_\theta \hat{\beta}_i = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Large-sample: ignore X_i-dependence in $\hat{\beta}$, & thus θ-dependence in X_i

$$\partial_{X_i}\hat{\beta}\,\partial_\theta X_i = \partial_{X_i}\left((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}\right)\partial_\theta X_i = 0$$

- Propagating through $\partial_{X_i}\hat{\beta}$ is as expensive as the main propagation of $\partial_\theta \hat{V}_{i+1}$
- Differentiating noise, $\partial_{X_i}\hat{\beta} = \partial_{X_i}(\beta_\infty - (\hat{\beta} - \beta_\infty)) = \partial_{X_i}(\hat{\beta} - \beta_\infty)) = \partial_{X_i}\epsilon_{\hat{\beta}}$
- Still important in presence of outliers/overfit, *eg*. in small samples

Coefficient Behavior and Dependencies in Large Samples

- Dependence upon θ gets propagated through the regression matrix

$$\partial_\theta \hat{\beta}_i = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Large-sample: ignore X_i-dependence in $\hat{\beta}$, & thus θ-dependence in X_i

$$\partial_{X_i}\hat{\beta}\,\partial_\theta X_i = \partial_{X_i}\left((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}\right)\partial_\theta X_i = 0$$

- Propagating through $\partial_{X_i}\hat{\beta}$ is as expensive as the main propagation of $\partial_\theta \hat{V}_{i+1}$
- Differentiating noise, $\partial_{X_i}\hat{\beta} = \partial_{X_i}(\beta_\infty - (\hat{\beta} - \beta_\infty)) = \partial_{X_i}(\hat{\beta} - \beta_\infty)) = \partial_{X_i}\epsilon_{\hat{\beta}}$
- Still important in presence of outliers/overfit, *eg*. in small samples

Coefficient Behavior and Dependencies in Large Samples

- Dependence upon θ gets propagated through the regression matrix

$$\partial_\theta \hat{\beta}_i = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Large-sample: ignore X_i-dependence in $\hat{\beta}$, & thus θ-dependence in X_i

$$\partial_{X_i}\hat{\beta}\,\partial_\theta X_i = \partial_{X_i}\left((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}\right)\partial_\theta X_i = 0$$

- Propagating through $\partial_{X_i}\hat{\beta}$ is as expensive as the main propagation of $\partial_\theta \hat{V}_{i+1}$
- Differentiating noise, $\partial_{X_i}\hat{\beta} = \partial_{X_i}(\beta_\infty - (\hat{\beta} - \beta_\infty)) = \partial_{X_i}(\hat{\beta} - \beta_\infty)) = \partial_{X_i}\epsilon_{\hat{\beta}}$
- Still important in presence of outliers/overfit, *eg*. in small samples

Coefficient Behavior and Dependencies in Large Samples

- Dependence upon θ gets propagated through the regression matrix

$$\partial_\theta \hat{\beta}_i = \left(\phi(X_i)'\phi(X_i)\right)^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Large-sample: ignore X_i-dependence in $\hat{\beta}$, & thus θ-dependence in X_i

$$\partial_{X_i}\hat{\beta}\,\partial_\theta X_i = \partial_{X_i}\left(\left(\phi(X_i)'\phi(X_i)\right)^{-1}\phi(X_i)'\hat{V}_{i+1}\right)\partial_\theta X_i = 0$$

- Propagating through $\partial_{X_i}\hat{\beta}$ is as expensive as the main propagation of $\partial_\theta \hat{V}_{i+1}$
- Differentiating noise, $\partial_{X_i}\hat{\beta} = \partial_{X_i}(\beta_\infty - (\hat{\beta} - \beta_\infty)) = \partial_{X_i}(\hat{\beta} - \beta_\infty)) = \partial_{X_i}\epsilon_{\hat{\beta}}$
- Still important in presence of outliers/overfit, *eg*. in small samples

Coefficient Behavior and Dependencies in Large Samples

- Dependence upon θ gets propagated through the regression matrix

$$\partial_\theta \hat{\beta}_i = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Large-sample: ignore X_i-dependence in $\hat{\beta}$, & thus θ-dependence in X_i

$$\partial_{X_i}\hat{\beta}\,\partial_\theta X_i = \partial_{X_i}\left((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}\right)\partial_\theta X_i = 0$$

- Propagating through $\partial_{X_i}\hat{\beta}$ is as expensive as the main propagation of $\partial_\theta \hat{V}_{i+1}$
- Differentiating noise, $\partial_{X_i}\hat{\beta} = \partial_{X_i}(\beta_\infty - (\hat{\beta} - \beta_\infty)) = \partial_{X_i}(\hat{\beta} - \beta_\infty)) = \partial_{X_i}\epsilon_{\hat{\beta}}$
- Still important in presence of outliers/overfit, *eg*. in small samples

Coefficient Behavior and Dependencies in Large Samples

- Dependence upon θ gets propagated through the regression matrix

$$\partial_\theta \hat{\beta}_i = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Large-sample: ignore X_i-dependence in $\hat{\beta}$, & thus θ-dependence in X_i

$$\partial_{X_i}\hat{\beta}\,\partial_\theta X_i = \partial_{X_i}\left((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\hat{V}_{i+1}\right)\partial_\theta X_i = 0$$

- Propagating through $\partial_{X_i}\hat{\beta}$ is as expensive as the main propagation of $\partial_\theta \hat{V}_{i+1}$
- Differentiating noise, $\partial_{X_i}\hat{\beta} = \partial_{X_i}(\beta_\infty - (\hat{\beta} - \beta_\infty)) = \partial_{X_i}(\hat{\beta} - \beta_\infty)) = \partial_{X_i}\epsilon_{\hat{\beta}}$
- Still important in presence of outliers/overfit, *eg*. in small samples

Coefficient Behavior and Dependencies in Large Samples

- Dependence upon θ gets propagated through the regression matrix

$$\partial_\theta \hat{\beta}_i = (\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\partial_\theta \hat{V}_{i+1}$$

- Large-sample: ignore X_i-dependence in $\hat{\beta}$, & thus θ-dependence in X_i

$$\partial_{X_i}\hat{\beta}\,\partial_\theta X_i = \partial_{X_i}\left((\phi(X_i)'\phi(X_i))^{-1}\phi(X_i)'\,\hat{V}_{i+1}\right)\partial_\theta X_i = 0$$

- Propagating through $\partial_{X_i}\hat{\beta}$ is as expensive as the main propagation of $\partial_\theta \hat{V}_{i+1}$
- Differentiating noise, $\partial_{X_i}\hat{\beta} = \partial_{X_i}(\beta_\infty - (\hat{\beta} - \beta_\infty)) = \partial_{X_i}(\hat{\beta} - \beta_\infty)) = \partial_{X_i}\epsilon_{\hat{\beta}}$
- Still important in presence of outliers/overfit, *eg*. in small samples

AD-on-LSMC Accuracy: Large-Sample Propagation

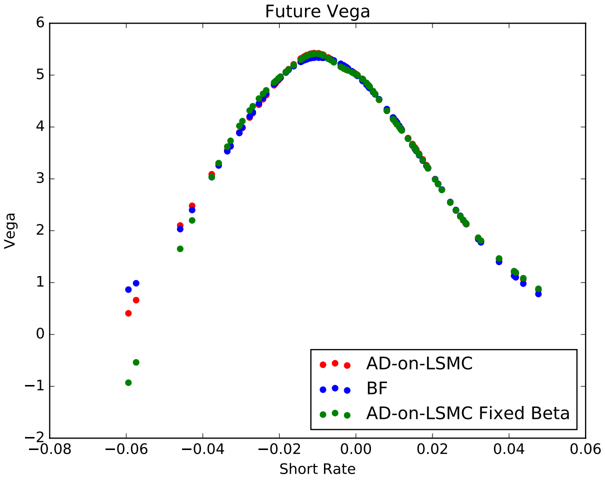

Figure: AD-on-LSMC Vegas with no $\partial_{X_i}\hat{\beta}$ propagation *vs.* Brute-Force: 10-into-16 Bermudan at 5Y Observation

AD-on-LSMC Accuracy: Large-Sample Propagation

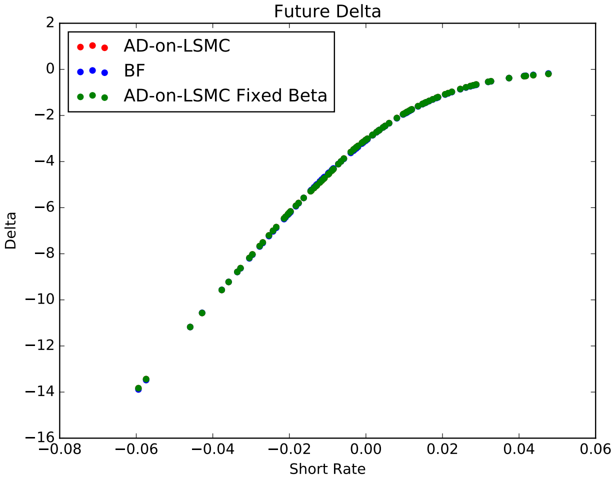

Figure: AD-on-LSMC Deltas with no $\partial_{X_i}\hat{\beta}$ propagation *vs.* Brute-Force: 10-into-16 Bermudan at 5Y Observation

AD-on-LSMC: Propagation Mode

- AD evaluates chain rule in either tangent (forward) or adjoint (reverse) modes
- Tangent costs ($\approx$) $\mathcal{O}(N_{ins})$ while adjoint costs ($\approx$) $\mathcal{O}(N_{outs})$

$$\text{CVA}: \; N_{ins} = N_\theta \; \& \; N_{outs} = 1 \implies \text{adjoint}$$

$$\text{MVA}: \; N_{ins} = N_\theta \; \& \; N_{outs} = N_T \cdot N_P \implies \text{tangent}$$

- MVA is *not* a Greek: Greeks over *all* exposures, $\partial_\theta \hat{V}_{p,i}$, are *inputs*

Future Greeks for CVA Greeks and MVA (Appendix)

- Mild difference between future Greeks for CVA, and future Greeks for MVA
- Future Greeks for CVA include trajectory: requires additional propagation

$$\partial_\theta \mathrm{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)}\, \partial_\theta V(t)\, dt\right]$$

- Future Greeks for MVA are along a fixed trajectory: no additional propagation

$$\mathrm{MVA} = \mathbb{E}_0\left[\int_0^T \mathrm{IM}(\partial_\theta V(t))\, dt\right]$$

Future Greeks for CVA Greeks and MVA

- Mild difference between future Greeks for CVA, and future Greeks for MVA
- Future Greeks for CVA include trajectory: requires additional propagation

$$\partial_\theta \mathrm{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)}\, \partial_\theta V(t)\, dt\right]$$

- Future Greeks for MVA are along a fixed trajectory: no additional propagation

$$\mathrm{MVA} = \mathbb{E}_0\left[\int_0^T \mathrm{IM}(\partial_\theta V(t))\, dt\right]$$

Future Greeks for CVA Greeks and MVA

- Mild difference between future Greeks for CVA, and future Greeks for MVA
- Future Greeks for CVA include trajectory: requires additional propagation

$$\partial_\theta \mathrm{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)}\, \partial_\theta V(t, X(t,\theta), \theta)\, dt\right]$$

- Future Greeks for MVA are along a fixed trajectory: no additional propagation

$$\mathrm{MVA} = \mathbb{E}_0\left[\int_0^T \mathrm{IM}(\partial_\theta V(t))\, dt\right]$$

Future Greeks for CVA Greeks and MVA

- Mild difference between future Greeks for CVA, and future Greeks for MVA
- Future Greeks for CVA include trajectory: requires additional propagation

$$\partial_\theta \mathrm{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)} \left(\partial_\theta V(t) + \partial_{X(t)} V(t) \cdot \partial_\theta X(t)\right) dt\right]$$

- Future Greeks for MVA are along a fixed trajectory: no additional propagation

$$\mathrm{MVA} = \mathbb{E}_0\left[\int_0^T \mathrm{IM}(\partial_\theta V(t))\, dt\right]$$

Future Greeks for CVA Greeks and MVA

- Mild difference between future Greeks for CVA, and future Greeks for MVA
- Future Greeks for CVA include trajectory: requires additional propagation

$$\partial_\theta \mathrm{CVA} = \mathbb{E}_0 \left[\int_0^T 1_{(V(t)>0)} \left(\partial_\theta V(t) + \partial_{X(t)} V(t) \cdot \partial_\theta X(t) \right) dt \right]$$

- Future Greeks for MVA are along a fixed trajectory: no additional propagation

$$\mathrm{MVA} = \mathbb{E}_0 \left[\int_0^T \mathrm{IM}(\partial_\theta V(t)) \, dt \right]$$

Future Greeks for CVA Greeks and MVA

- Mild difference between future Greeks for CVA, and future Greeks for MVA
- Future Greeks for CVA include trajectory: requires additional propagation

$$\partial_\theta \text{CVA} = \mathbb{E}_0 \left[\int_0^T 1_{(V(t)>0)} \left(\partial_\theta V(t) + \partial_{X(t)} V(t) \cdot \partial_\theta X(t) \right) dt \right]$$

- Future Greeks for MVA are along a fixed trajectory: no additional propagation

$$\text{MVA} = \mathbb{E}_0 \left[\int_0^T \text{IM}(\partial_\theta V(t))\, dt \right]$$

numerix

Future Greeks for CVA Greeks and MVA

- Mild difference between future Greeks for CVA, and future Greeks for MVA
- Future Greeks for CVA include trajectory: requires additional propagation

$$\partial_\theta \mathrm{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)}\left(\partial_\theta V(t) + \partial_{X(t)} V(t)\cdot\partial_\theta X(t)\right) dt\right]$$

- Future Greeks for MVA are along a fixed trajectory: no additional propagation

$$\mathrm{MVA} = \mathbb{E}_0\left[\int_0^T \mathrm{IM}(\partial_\theta V(t, X(t,\theta),\theta))\, dt\right]$$

Future Greeks for CVA Greeks and MVA

- Mild difference between future Greeks for CVA, and future Greeks for MVA
- Future Greeks for CVA include trajectory: requires additional propagation

$$\partial_\theta \mathrm{CVA} = \mathbb{E}_0\left[\int_0^T 1_{(V(t)>0)}\left(\partial_\theta V(t) + \partial_{X(t)} V(t)\cdot \partial_\theta X(t)\right) dt\right]$$

- Future Greeks for MVA are along a fixed trajectory: no additional propagation

$$\mathrm{MVA} = \mathbb{E}_0\left[\int_0^T \mathrm{IM}(\partial_\theta V(t, X(t,\theta), \theta)|_{X(t)})\, dt\right]$$

Future Greeks for CVA Greeks and MVA

- Mild difference between future Greeks for CVA, and future Greeks for MVA
- Future Greeks for CVA include trajectory: requires additional propagation

$$\partial_\theta \mathrm{CVA} = \mathbb{E}_0 \left[\int_0^T 1_{(V(t)>0)} \left(\partial_\theta V(t) + \partial_{X(t)} V(t) \cdot \partial_\theta X(t) \right) dt \right]$$

- Future Greeks for MVA are along a fixed trajectory: no additional propagation

$$\mathrm{MVA} = \mathbb{E}_0 \left[\int_0^T \mathrm{IM}(\partial_\theta V(t))\, dt \right]$$

MVA: Motivation and Logistics 1 (Appendix)

- MVA is lifetime funding cost of IM, and IM is sensitivity-based VaR[4]

$$\text{MVA} = \mathbb{E}_0\left[\int_0^T \text{IM}(\partial_{Q(t)} V(t))\, dt\right]$$

- IM is additional collateral to mitigate counterparty risk over MPoR ($\sim$ 10D)
- Bilateral IM: both c/parties post to 3^{rd}-party custodians $\implies$ needs funding
- In practice, portfolio hedges attract bilateral &/or clearing-house IM too
- MVA reflects funding costs in valuations $\implies$ spectre of FVA debate

[4] See Green and Kenyon ('15) for detailed derivation

Swap IM Projections

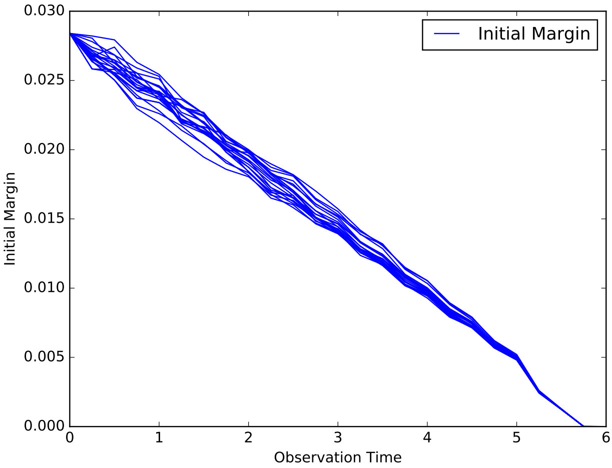

Figure: Delta-IM for a vanilla swap: just applying SIMM rule, not CCH rule

Swaption IM Projections

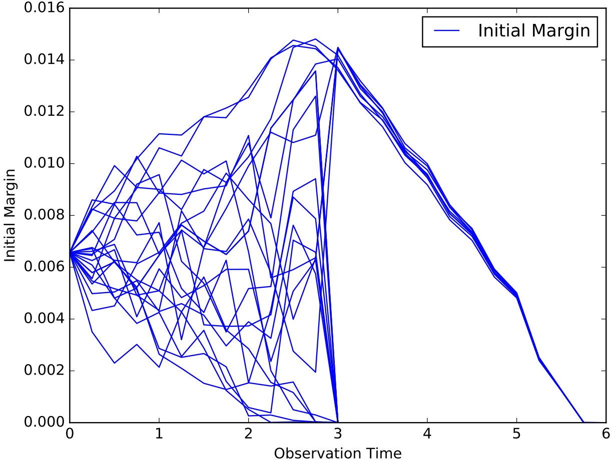

Figure: Delta-IM for a swaption

Bermudan IM Projections

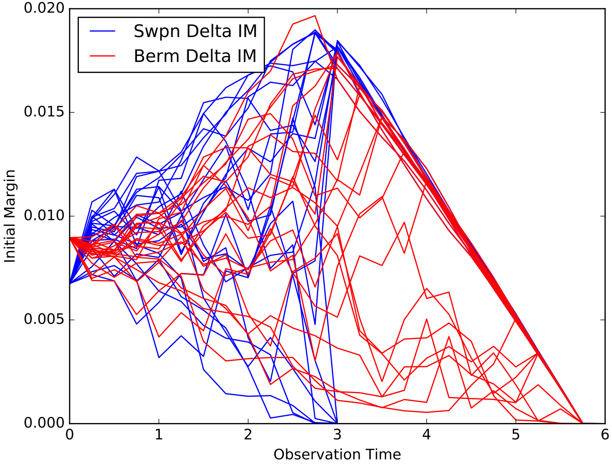

Figure: Delta-IM for a Bermudan