Building Closed-Form Formula for Real-Time Derivative Pricing and Greeks Calculation Using Offline Simulation

SHEN Haihui（沈海辉）

Sino-US Global Logistics Institute
Shanghai Jiao Tong University

Joint work with Guangxin Jiang (SHU), and Jeff Hong (Fudan)

@ FERM2019, SUFE

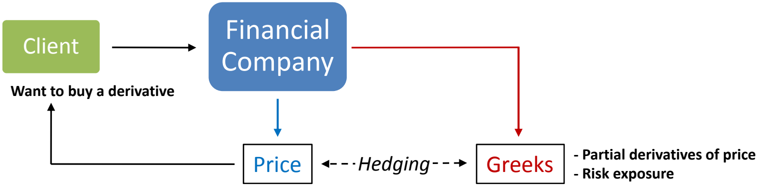
August 30, 2019

上海交通大學 Shanghai Jiao Tong University | 董浩云航运与物流研究院 CY TUNG Institute of Maritime and Logistics
中美物流研究院 Sino-US Global Logistics Institute

Outline

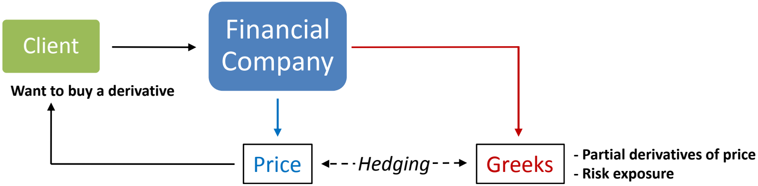
Financial Background

- A financial company may need to immediately quote the price of a derivative upon enquiry, and know the Greeks for hedging (once the transaction is made).



Financial Background

- A financial company may need to immediately quote the price of a derivative upon enquiry, and know the Greeks for hedging (once the transaction is made).



- Fair price and Greeks change rapidly as the market conditions change ⇒ Real-time problem.
  - For simple model, analytical formulae of the derivative price and Greeks are available (no difficulty in real-time use).
  - For realistic model, Monte Carlo simulation is often required to estimate price and Greeks (cannot output results in real time).

Financial Background

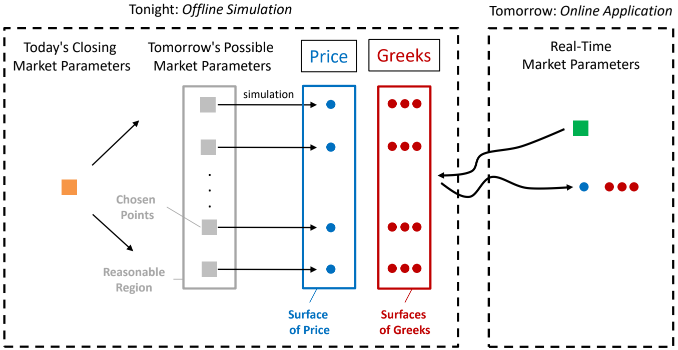
- Important Facts:
  1. Financial markets only open during the working hours.
  2. Market conditions tomorrow normally vary within some ranges from the closing conditions today.

Financial Background

- Important Facts:
  1. Financial markets only open during the working hours.
  2. Market conditions tomorrow normally vary within some ranges from the closing conditions today.
- How to take advantages of such facts?

Offline Simulation Online Application

- Offline Simulation: When market closes, run simulation to learn the surfaces of price and Greeks (over certain ranges of market parameters).
- Online Application: When market opens tomorrow, use the learned surfaces to quote real-time price and hedge risk.



Key Research Question

- How to construct surfaces of price and Greeks during offline simulation period?
  1. So that they can be used in a way like analytical formulae.
  2. So that they are accurate enough.
  3. So that the used price and Greeks are consistent.

Key Research Question

- How to construct surfaces of price and Greeks during offline simulation period?
  1. So that they can be used in a way like analytical formulae.
  2. So that they are accurate enough.
  3. So that the used price and Greeks are consistent.
- Consistency is defined as:

$$\widehat{G}^k(\mathbf{x}) = \frac{\partial \widehat{V}(\mathbf{x})}{\partial x_k},$$

  with the following notations:
  - $\mathbf{x} := (x_1, x_2, \ldots)^\intercal$ denotes the market parameters (factors);
  - $V(\mathbf{x})$ denotes the price of a derivative (or a portfolio);
  - $G^k(\mathbf{x}) := \partial V(\mathbf{x}) / \partial x_k$ denotes the Greeks;
  - $\widehat{V}(\mathbf{x})$ denotes the estimator of the price;
  - $\widehat{G}^k(\mathbf{x})$ denotes the estimator of Greeks.

Why Consistency Matters?

- Consistency between used price and Greeks is critical to
  - ensure effective hedging (P&L close to zero);
  - maintain stable balance sheet in accounting.

Why Consistency Matters?

- Consistency between used price and Greeks is critical to
  - ensure effective hedging (P&L close to zero);
  - maintain stable balance sheet in accounting.
- Our Theorem 1 shows that the fluctuation (variance) of the company's P&L will be smaller when consistency exits.

Why Consistency Matters?

- Consistency between used price and Greeks is critical to
  - ensure effective hedging (P&L close to zero);
  - maintain stable balance sheet in accounting.
- Our Theorem 1 shows that the fluctuation (variance) of the company's P&L will be smaller when consistency exits.
- Our Theorem 2 shows that to achieve the same hedging effect, the hedging cost in consistency case will be less than that in inconsistency case.

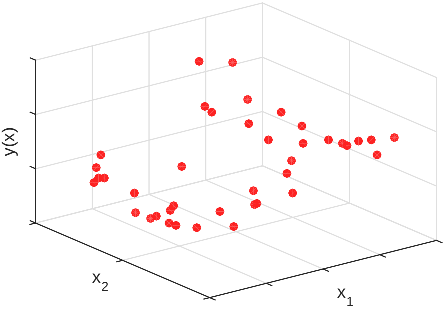
Stochastic Kriging

- Kriging, named after the South African mining engineer Danie G. Krige, is a method of interpolation.
  - Originally used to interpolate the altitude of a landscape.

Stochastic Kriging

- Kriging, named after the South African mining engineer Danie G. Krige, is a method of interpolation.
  - Originally used to interpolate the altitude of a landscape.

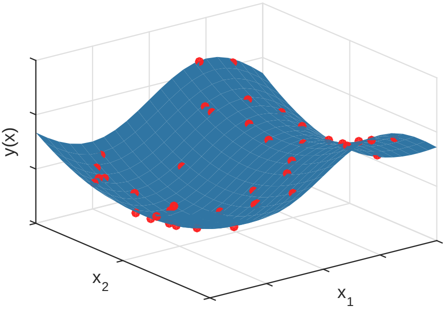
- Observing $(\mathbf{x}_1, y(\mathbf{x}_1)), (\mathbf{x}_2, y(\mathbf{x}_2)), \ldots$, we wish to predict $y(\mathbf{x}_0)$ for any $\mathbf{x}_0$.



Stochastic Kriging

- Kriging, named after the South African mining engineer Danie G. Krige, is a method of interpolation.
  - Originally used to interpolate the altitude of a landscape.

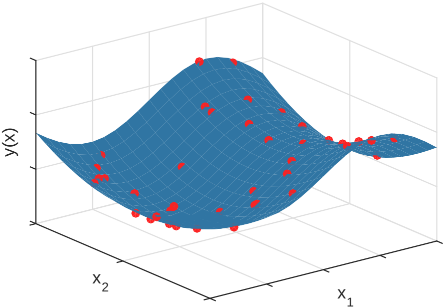
- Observing $(\mathbf{x}_1, y(\mathbf{x}_1)), (\mathbf{x}_2, y(\mathbf{x}_2)), \ldots$, we wish to predict $y(\mathbf{x}_0)$ for any $\mathbf{x}_0$.



Stochastic Kriging

- Kriging, named after the South African mining engineer Danie G. Krige, is a method of interpolation.
  - Originally used to interpolate the altitude of a landscape.

- Observing $(\mathbf{x}_1, y(\mathbf{x}_1)), (\mathbf{x}_2, y(\mathbf{x}_2)), \ldots$, we wish to predict $y(\mathbf{x}_0)$ for any $\mathbf{x}_0$.



- Stochastic Kriging (SK) allows observation errors in y.
  - It is no longer exact interpolation.

Stochastic Kriging

- Surfaces constructed from SK possess analytical forms (linear combination of observed $y(\mathbf{x}_i)$).
  - $\surd$ Requirement 1: Used in a way like analytical formulae.

Stochastic Kriging

- Surfaces constructed from SK possess analytical forms (linear combination of observed $y(\mathbf{x}_i)$).
  - $\surd$ Requirement 1: Used in a way like analytical formulae.

- Naturally, one may consider the following approaches.

Stochastic Kriging

- Surfaces constructed from SK possess analytical forms (linear combination of observed $y(\mathbf{x}_i)$).
  - $\surd$ Requirement 1: Used in a way like analytical formulae.
- Naturally, one may consider the following approaches.
- Approach A: Construct surfaces for price and Greeks using SK, *separately*.
  - $\surd$ Requirement 2: Accuracy.
  - $\times$ Requirement 3: Consistency.

Stochastic Kriging

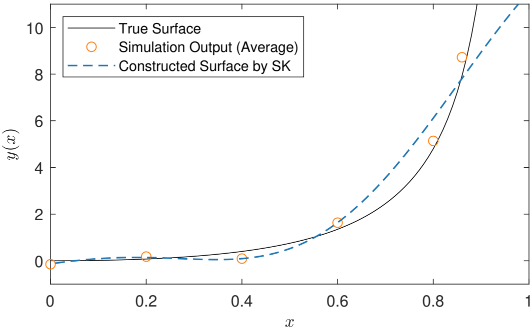
- Surfaces constructed from SK possess analytical forms (linear combination of observed $y(\mathbf{x}_i)$).
  - √ Requirement 1: Used in a way like analytical formulae.
- Naturally, one may consider the following approaches.
- Approach A: Construct surfaces for price and Greeks using SK, *separately*.
  - √ Requirement 2: Accuracy.
  - × Requirement 3: Consistency.
- Approach B: Construct price surface using SK, and get Greeks by differentiating the price surface.
  - × Requirement 2: Accuracy.
  - √ Requirement 3: Consistency.

Gradient Enhanced Stochastic Kriging

- Gradient Enhanced Stochastic Kriging (GESK), also known as co-kriging, *combines the observations of price and Greeks together to construct the price surface.*

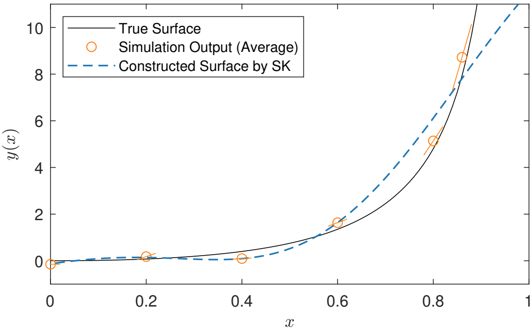
Gradient Enhanced Stochastic Kriging

- Gradient Enhanced Stochastic Kriging (GESK), also known as co-kriging, *combines the observations of price and Greeks together to construct the price surface.*



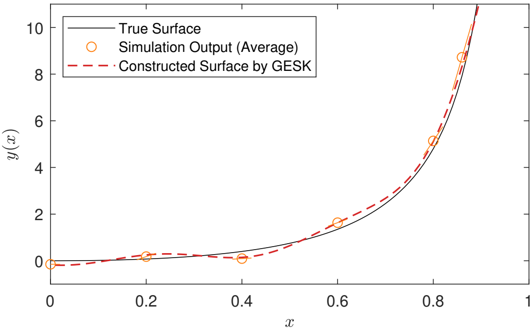
Gradient Enhanced Stochastic Kriging

- Gradient Enhanced Stochastic Kriging (GESK), also known as co-kriging, *combines the observations of price and Greeks together to construct the price surface.*



Gradient Enhanced Stochastic Kriging

- Gradient Enhanced Stochastic Kriging (GESK), also known as co-kriging, *combines the observations of price and Greeks together to construct the price surface.*



Gradient Enhanced Stochastic Kriging

- Recall that
  - Approach A: Construct surfaces for price and Greeks using SK, *separately*.
    - √ Requirement 2: Accuracy.
    - × Requirement 3: Consistency.
  - Approach B: Construct price surface using SK, and get Greeks by differentiating the price surface.
    - × Requirement 2: Accuracy.
    - √ Requirement 3: Consistency.

Gradient Enhanced Stochastic Kriging

- Recall that
  - Approach A: Construct surfaces for price and Greeks using SK, *separately*.
    - ✓ Requirement 2: Accuracy.
    - × Requirement 3: Consistency.
  - Approach B: Construct price surface using SK, and get Greeks by differentiating the price surface.
    - × Requirement 2: Accuracy.
    - ✓ Requirement 3: Consistency.
- Approach C: Construct price surface using GESK, and get Greeks by differentiating the price surface.
  - ?? Requirement 2: Accuracy.
  - ✓ Requirement 3: Consistency.

Gradient Enhanced Stochastic Kriging

- Recall that
  - Approach A: Construct surfaces for price and Greeks using SK, *separately*.
    - √ Requirement 2: Accuracy.
    - × Requirement 3: Consistency.
  - Approach B: Construct price surface using SK, and get Greeks by differentiating the price surface.
    - × Requirement 2: Accuracy.
    - √ Requirement 3: Consistency.
- Approach C: Construct price surface using GESK, and get Greeks by differentiating the price surface.
  - √√ Requirement 2: Accuracy.
  - √ Requirement 3: Consistency.
- Our Theorem 4 shows that, the accuracy of **price and Greeks** in Approach C are **higher** than that in Approach A.

Black-Scholes Model

- Consider a European call option (maturity $T = 1$ year, strike price $K = 105$).
  - The underlying stock price is driven by a geometric Brownian motion.
  - Option price formula is given by B-S formula, and Greeks formulae can be obtained by differentiating the price formula.

Black-Scholes Model

- Consider a European call option (maturity $T = 1$ year, strike price $K = 105$).
  - The underlying stock price is driven by a geometric Brownian motion.
  - Option price formula is given by B-S formula, and Greeks formulae can be obtained by differentiating the price formula.
- Pretend that the formulae are unknown and compare Approaches A, B, C under OSOA.

Black-Scholes Model

- Consider a European call option (maturity $T = 1$ year, strike price $K = 105$).
  - The underlying stock price is driven by a geometric Brownian motion.
  - Option price formula is given by B-S formula, and Greeks formulae can be obtained by differentiating the price formula.
- Pretend that the formulae are unknown and compare Approaches A, B, C under OSOA.
- Closing market parameters: stock price $S_t = 100$, volatility $\sigma = 0.2$, interest rate $r = 0.02$.

Black-Scholes Model

- Consider a European call option (maturity $T = 1$ year, strike price $K = 105$).
  - The underlying stock price is driven by a geometric Brownian motion.
  - Option price formula is given by B-S formula, and Greeks formulae can be obtained by differentiating the price formula.
- Pretend that the formulae are unknown and compare Approaches A, B, C under OSOA.
- Closing market parameters: stock price $S_t = 100$, volatility $\sigma = 0.2$, interest rate $r = 0.02$.
- Sample 20 design points

$$\mathbf{x}_i = (S_t, \sigma, r) \in [80, 120] \times [0.01, 0.3] \times [0.001, 0.1]$$

using Latin hypercube sampling method.

Surface Accuracy

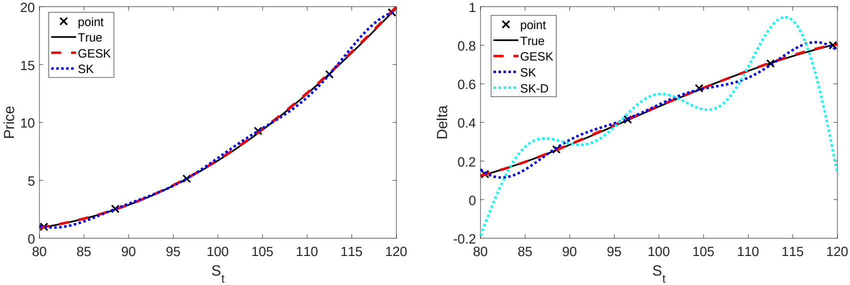


Figure: Price (left) and `delta` (right) surfaces for $S_t \in [80, 120]$.

Surface Accuracy

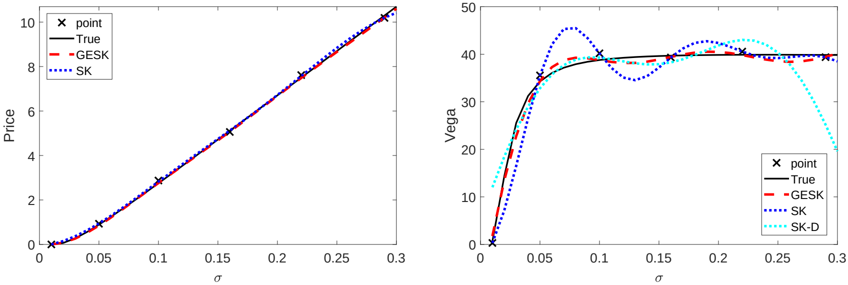


Figure: Price (left) and vega (right) surfaces for $\sigma \in [0.01, 0.3]$.

Surface Accuracy

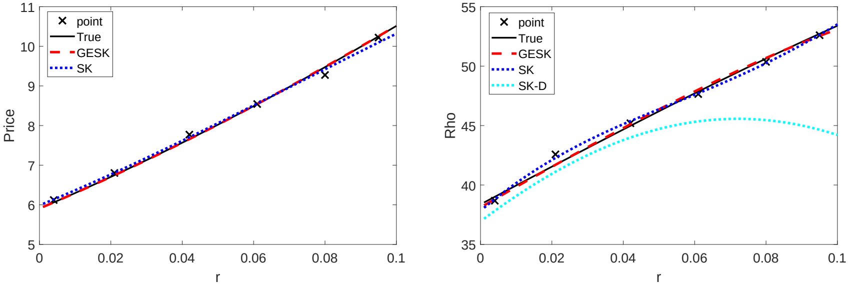


Figure: Price (left) and `rho` (right) surfaces for $r \in [0.001, 0.1]$.

Delta Hedging Effect

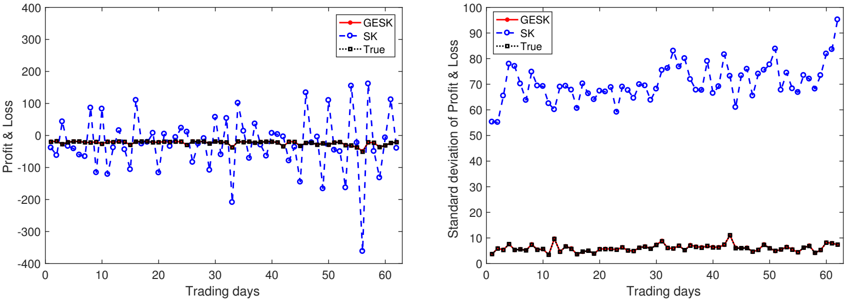


Figure: P&L under one specific stock path (left) and standard deviation over 100 stock paths (right).

Variance Gamma Model

- Consider a portfolio with 5 Asian options and 5 lookback options, based on 5 stocks.
  - (1) Apple, (2) Facebook, (3) Netflix, (4) Alibaba, (5) Tesla.
  - The stock price is modeled by the exponential variance gamma process.
  - Analytical formulae of price and Greeks are unavailable.

Variance Gamma Model

- Consider a portfolio with 5 Asian options and 5 lookback options, based on 5 stocks.
  - (1) Apple, (2) Facebook, (3) Netflix, (4) Alibaba, (5) Tesla.
  - The stock price is modeled by the exponential variance gamma process.
  - Analytical formulae of price and Greeks are unavailable.

- Based on the data from Yahoo Finance on 9th November 2018, we set the closing stock price S and yield, and calibrate parameters (σ, ν, θ).

	Apple, Inc.	Facebook, Inc.	Netflix, Inc.	Alibaba	Tesla, Inc.
S	204.47	144.96	303.47	144.85	350.51
yield	1.21%	0	0	0	0
σ	0.2636	0.2625	0.4012	0.2842	0.4660
ν	0.0387	0.0355	0.0394	0.0017	0.0933
θ	-0.5185	-0.8288	-1.2344	-2.6984	-1.1459

Delta Hedging Effect

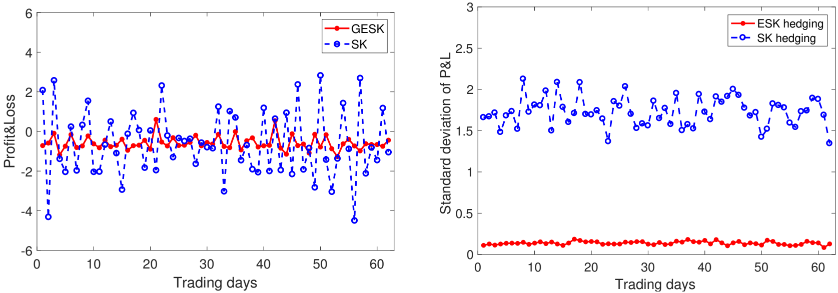


Figure: P&L under one specific stock path (left) and standard deviation over 50 stock paths (right).

Conclusion

- Under the perspective of offline simulation online application, simulation can be used to solve real-time problem, e.g., real-time pricing and hedging.
- For the pricing and hedging problem, consistency between price and Greeks matters.
- Price and Greeks surfaces constructed using GESK are consistent and accurate, which yield satisfactory hedging effect.

Thank You!

SHEN Haihui
shenhaihui@sjtu.edu.cn

Supplement

Hedging Effect

- Use *delta hedging* as an example.
- Assume there is only one underlying asset and only the asset price S_t is changing (other market factors in $\mathbf{x}$ keep unchanged).
- Profit & Loss of the hedged derivative (portfolio).
  - Consistency:

$$L(S_t) = -[V(S_t) - V(s_0)] + \Delta(s_0)[S_t - s_0], \ \Delta(s_0) = V'(s_0)$$

  - Inconsistency:

$$L^{\dagger}(S_t) = -[V(S_t) - V(s_0)] + \Delta^{\dagger}(s_0)[S_t - s_0], \ \Delta^{\dagger}(s_0) \neq V'(s_0)$$

Hedging Effect

Theorem (1)

Suppose that the underlying asset S_t is driven by the exponential family of stochastic process $S_t = s_0 \exp(at + \sqrt{t}X_t)$, where $X_t \xrightarrow{d} X$ as $t \to 0^+$ with $\mathbb{E}\left[X^4\right] < \infty$ and $\mathrm{Var}[X] > 0$. Moreover, assume that there exist $h > 0$ and $t_h > 0$ such that $\sup_{0<t\leq t_h} \mathbb{E}\left[e^{\theta X_t}\right] < \infty$ for all $|\theta| \leq h$, and the second derivative of $V(s)$ is bounded above. Then for the P&L $L(S_t)$ and $L^\dagger(S_t)$ defined above, there exists $\tau > 0$ such that $\mathrm{Var}[L(S_t)] < \mathrm{Var}[L^\dagger(S_t)]$ for $t < \tau$.

Remark

If one is willing to assumes that $V(s_0)$ is approximately linear over a small range around s_0 (i.e., perfect hedging), then the result of Theorem 1 can be obtained without assuming any form of S_t. Indeed, in this case, $L(S_t) \approx 0$ while $L^\dagger(S_t) \neq 0$ for small t.

Hedging Cost

- One may consider the problem the other way around, that is, to achieve the same hedging effect, what is the difference in efforts when consistency does and does not exist?
- Hedging cost:
  - Consistency: $C = \left|\Delta(s_0) - \widetilde{\Delta}\right| s_0 d$
  - Inconsistency: $C^\dagger = \left|\Delta^\dagger(s_0) - \widetilde{\Delta}\right| s_0 d + \sum_{i=1}^m \left|\Delta_i^\dagger - \Delta_{i-1}^\dagger\right| S_{t_i} d.$

Theorem (2)

Suppose in inconsistency case the risk manager needs to conduct a series of hedging at time $t_1, \ldots, t_m$ with $0 < t_1 < \cdots < t_m < t$, which successively adjusts the position to $\Delta_1^\dagger, \ldots, \Delta_m^\dagger$ such that $\Delta_m^\dagger = \Delta(s_0)$, for some $m \geq 1$, in order to achieve the the same hedging effect in the consistency case. Moreover, assume that $\mathbb{E}[S_{t_i}] = s_0$, for $i = 1, \ldots, m$. Then for the hedging cost C defined and $C^\dagger$ defined above, $C \leq \mathbb{E}[C^\dagger]$.

Stochastic Kriging

- Let $\mathbf{x} \in \Re^d$ be the market factors scenarios, and $\mathcal{Y}(\mathbf{x})$ be the derivative price.

$$\mathcal{Y}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^\top \boldsymbol{\beta} + \mathsf{M}(\mathbf{x}),$$

  where $\mathsf{M}(\mathbf{x})$ is a Gaussian random field with zero mean.
- $\mathcal{Y}(\mathbf{x})$ is observed with random noise,

$$Y_l(\mathbf{x}) = \mathcal{Y}(\mathbf{x}) + \varepsilon_l(\mathbf{x}) = \mathbf{f}(\mathbf{x})^\top \boldsymbol{\beta} + \mathsf{M}(\mathbf{x}) + \varepsilon_l(\mathbf{x}),$$

  where $\varepsilon_l(\mathbf{x})$ is the simulation error along the l-th sample path.
- Suppose that we have n design points $\mathbf{x}_i$, $i = 1, \ldots, n$, and on design point $\mathbf{x}_i$ the simulation is run for m_i replications:

$$\overline{Y}(\mathbf{x}_i) = \frac{1}{m_i} \sum_{l=1}^{m_i} Y_l(\mathbf{x}_i), \text{ and } \overline{\varepsilon}(\mathbf{x}_i) = \frac{1}{m_i} \sum_{l=1}^{m_i} \varepsilon_l(\mathbf{x}_i).$$

- The mean squared error (MSE) optimal predictor of $\mathcal{Y}(\mathbf{z})$ is given by

$$\hat{\mathcal{Y}}(\mathbf{z}) = \mathbf{f}(\mathbf{z})^\top \boldsymbol{\beta} + \gamma(\mathbf{z})^\top (\Gamma + \Sigma)^{-1} (\overline{\mathbf{Y}} - \mathbf{F}\beta).$$

Gradient Enhanced Stochastic Kriging

- Incorporates the response surface's gradient estimators into SK to improve the prediction accuracy of the response surface

$$\mathcal{D}^k(\mathbf{x}) = \frac{\partial}{\partial x_k}\mathcal{Y}(\mathbf{x}) = \left(\frac{\partial}{\partial x_k}\mathbf{f}(\mathbf{x})\right)^\top \boldsymbol{\beta} + \frac{\partial}{\partial x_k}\mathsf{M}(\mathbf{x}).$$

- The GESK models $D_l^k(\mathbf{x}), k = 1, \ldots, d$, as

$$D_l^k(\mathbf{x}) = \mathcal{D}^k(\mathbf{x}) + \epsilon_l^k(\mathbf{x}) = \left(\frac{\partial}{\partial x_k}\mathbf{f}(\mathbf{x})\right)^\top \boldsymbol{\beta} + \frac{\partial}{\partial x_k}\mathsf{M}(\mathbf{x}) + \epsilon_l^k(\mathbf{x}).$$

- $\mathcal{Y}(\mathbf{z})$ is predicted by

$$\widetilde{\mathcal{Y}}(\mathbf{z}) = \mathbf{f}(\mathbf{z})^\top \boldsymbol{\beta} + \gamma_+(\mathbf{z})^\top (\Gamma_+ + \Sigma_+)^{-1}(\overline{\mathbf{Y}}_+ - \mathbf{F}_+\boldsymbol{\beta}).$$

- $\mathcal{D}^k(\mathbf{z})$ is predicted by

$$\partial_k\widetilde{\mathcal{Y}}(\mathbf{z}) = (\partial_k\mathbf{f}(\mathbf{z}))^\top \boldsymbol{\beta} + (\partial_k\boldsymbol{\gamma}_+(\mathbf{z}))^\top (\Gamma_+ + \Sigma_+)^{-1}(\overline{\mathbf{Y}}_+ - \mathbf{F}_+\boldsymbol{\beta}).$$

Accuracy Analysis

Theorem (4)

Suppose that $\boldsymbol{\beta}$, τ^2, $\boldsymbol{\theta}$ and Σ_+ are known, then the MSE of $\widetilde{\mathcal{Y}}(\mathbf{z})$ is smaller than the MSE of $\hat{\mathcal{Y}}(\mathbf{z})$, i.e.,

$$\mathbb{E}\left[\left(\widetilde{\mathcal{Y}}(\mathbf{z}) - \mathcal{Y}(\mathbf{z})\right)^2\right] < \mathbb{E}\left[\left(\hat{\mathcal{Y}}(\mathbf{z}) - \mathcal{Y}(\mathbf{z})\right)^2\right],$$

and the MSE of $\frac{\partial}{\partial z_k}\widetilde{\mathcal{Y}}(\mathbf{z})$ is smaller than the MSE of $\hat{\mathcal{D}}^k(\mathbf{z})$, i.e.,

$$\mathbb{E}\left[\left(\frac{\partial}{\partial z_k}\widetilde{\mathcal{Y}}(\mathbf{z}) - \mathcal{D}^k(\mathbf{z})\right)^2\right] < \mathbb{E}\left[\left(\hat{\mathcal{D}}^k(\mathbf{z}) - \mathcal{D}^k(\mathbf{z})\right)^2\right],$$

for $k = 1, \ldots, d$.

Proposition (1)

$\frac{\partial}{\partial z_k}\widetilde{\mathcal{Y}}(\mathbf{z}) = \widetilde{\mathcal{D}}^k(\mathbf{z})$, *for all $\mathbf{z}$ and $k = 1, \ldots, d$.*

Surface Accuracy for B-S Model

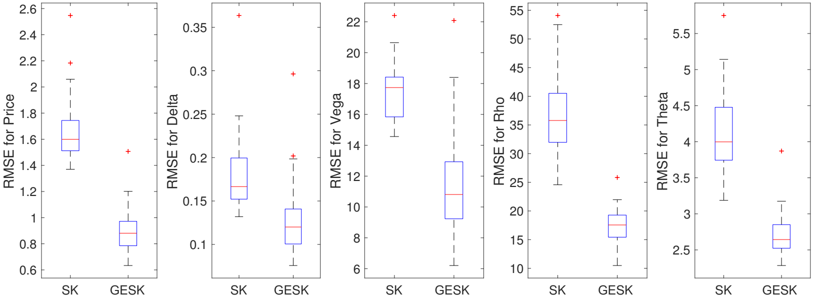


Figure: Boxplots of RMSE for price, `delta`, `vega`, `rho`, and `theta` surfaces.

Surface Accuracy for Variance Gamma Model