Project 1: Pricing with linear regression

(First discussion: Oct 4; Last questions: Oct 25; Deadline: Nov 1)

Consider a money market account offering a constant riskless interest rate r and four risky assets with prices following the dynamics

$$S_t^j = s_0^j \exp\left(\left[\mu_j - \frac{1}{2}\sigma_j^2\right] t + \sigma_j W_t^j\right), \quad t \geq 0,\ j = 1, \dots, 4,$$

for parameters s_0^j, μ_j, σ_j and independent standard Brownian motions W^j, $j = 1, \dots, 4$.

1. Determine the time-0 value V of a derivatives portfolio consisting of two European call options with time-T payoff $(S_T^j - K)^+$, $j = 1, 2$, and two European put options with time-T payoff $(K - S_T^j)^+$, $j = 3, 4$, with common maturity $T > 0$ and strike price $K > 0$.

   Hint: One has $V = \sum_{j=1}^4 V_j$, where V_j is the price of the j-th option given by the Black–Scholes formula for a European call or put option.

2. Now, denote by $V(x)$, $x \in [80, 120]^4$, the time-0 value of the derivatives portfolio if $s_0 = x$, $K = 100$, $T = 1$, $r = 2\%$ and $\sigma_j = (j + 1)10\%$, $j = 1, \dots, 4$.

   Simulate $x^i \in \mathbb{R}^4$, $i = 1, \dots, 1000$, independently from the uniform distribution over $[80, 120]^4$ and let $y_i = V(x^i)$, $i = 1, \dots, 1000$.

   Approximate $V(x)$, $x \in [80, 120]^4$, with an affine function of the form $f(x) = b_0 + \sum_{j=1}^4 b_j x_j$ by solving the normal equation $A^T A b = A^T y$ for

$$A = \begin{pmatrix} 1 & x_1^1 & \dots & x_4^1 \\ 1 & x_1^2 & \dots & x_4^2 \\ \vdots & \vdots & \dots & \vdots \\ 1 & x_1^{1000} & \dots & x_4^{1000} \end{pmatrix}, \quad b = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_4 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ \vdots \\ y_{1000} \end{pmatrix}.$$

   (i) Report the approximation error $f(x) - V(x)$ for $x = (100, 100, 100, 100)$;

   (ii) Estimate the L^1 approximation error

$$\|f - V\|_1 = \frac{1}{40^4} \int_{[80,120]^4} |f(x) - V(x)| dx \ \text{ and}$$

   on a simulated test set $\tilde{x}^i \in \mathbb{R}^4$, $i = 1, \dots, n$, and derive an approximate 95% confidence interval for $\|f - V\|_1$ from the sample standard deviation. Choose the size n of the test set so that the width of the approximate 95% confidence interval does not exceed the estimate of $\|f - V\|_1$.

3. Now, we want to approximate $V(x)$, $x \in [80, 120]^4$, with a second order polynomial of the form

$$g(x) = b_0 + \sum_{j=1}^4 b_j x_j + \sum_{1 \leq j \leq k \leq 4} b_{jk} x_j x_k.$$

a) Try to solve the extended normal equation $B^T Bb = B^T y$, where B is the extended design matrix corresponding to the features x_j, $1 \leq j \leq 4$, and $x_j x_k$, $1 \leq j \leq k \leq 4$, and $b = (b_0, b_j, b_{jk})$, $1 \leq j \leq k \leq 4$. What is the problem?

b) Determine the eigenvalues of the matrix $B^T B$.

c) Compute an approximate solution b to the equation $Bb = y$ by truncated pseudoinversion of B. Use 5-fold cross-validation to find an optimal truncation threshold.

d) Compute an approximate solution b to the equation $Bb = y$ with ridge regression. Use 5-fold cross-validation to find an optimal penalty parameter.

In the cases c)–d), report the approximation errors (i)–(ii) from above on a simulated test set $\tilde{x}^i$, $i = 1, \ldots, n$, and provide approximate 95% confidence intervals for (ii). Justify your choice of the size n of the test set.

4. Compute the time-0 value V_{mc} of a max-call option with time-T payoff $(\max_{1 \leq j \leq 4} S_T^j - K)^+$ for the case that $s_0 = (100, 100, 100, 100)$, $K = 100$, $T = 1$, $r = 2\%$ and $\sigma_j = (j+1)10\%$, $j = 1, \ldots, 4$.

   Hint: *V_{mc} is given by $e^{-rT}\mathbb{E}^{\mathbb{Q}}\left(\max_{1 \leq j \leq 4} S_T^j - K\right)^+$ for a measure $\mathbb{Q}$ under which the dynamics of the prices of the risky assets are*

$$S_t^j = s_0^j \exp\left(\left[r - \frac{1}{2}\sigma_j^2\right] t + \sigma_j \tilde{W}_t^j\right), \quad t \geq 0, \ j = 1, \ldots, 4,$$

   where $\tilde{W}^j$, $j = 1, \ldots, 4$, are independent standard Brownian motions under $\mathbb{Q}$. Therefore, V_{mc} can be approximated with a Monte Carlo average. Use at least 10,000 simulations to compute the Monte Carlo average.

5. Now, denote by $V_{\mathrm{mc}}(x)$, $x \in [80, 120]^4$, the (Monte-Carlo approximation of the) time-0 value of the max-call option if $s_0 = x$, $K = 100$, $T = 1$, $r = 2\%$ and $\sigma_j = (j+1)10\%$, $j = 1, \ldots, 4$.

   Simulate $x^i \in \mathbb{R}^4$, $i = 1, \ldots, 1000$, independently from the uniform distribution over $[80, 120]^4$ and let $y_i = V_{\mathrm{mc}}(x^i)$, $i = 1, \ldots, 1000$.

   Approximate $V_{\mathrm{mc}}(x)$, $x \in [80, 120]^4$, with a second order polynomial of the form

$$g(x) = b_0 + \sum_{j=1}^{4} b_j x_j + \sum_{1 \leq j \leq k \leq 4} b_{jk} x_j x_k$$

   with the methods c) and d) from 3.

   Report the approximation error $g(x) - V_{\mathrm{mc}}(x)$ for $x = (100, 100, 100, 100)$.