

The Finite Element Method for Option Pricing under Heston's Model

Final Report for

MA6621 Programming and Computing for Finance and Actuarial Science

HU Wei (SID: 54206763)

YAO Wuguannan (SID: 54266563)

HUANG Jiaheng (SID: 54304001)

Examiner:

Professor Benny Y.C. Hon

2016.05.10

Abstract

Option is one of the most important derivatives in financial markets. Since for some complicated types of options there are no available analytical solutions, we are devoted to applying Finite Element Method (FEM) for option pricing problem in this report. First, we try to resolve the problem of plain vanilla option pricing with analysis of accuracy of FEM by comparing with the analytical solution. Further, in order to demonstrate advantages of the FEM, we take a deeper analysis and derive some equations under Heston's model, which pushes forward the PDE formulation of the model into two-dimensional problem. Finally, we provide a set of examples for pricing European type option under Heston's model, and associated numerical results to demonstrate the accuracy, convergence and efficiency.

Contents

1.	Introduction	1
2.	Theory of Finite Element Method	3
2.1.	Variational Formulation	3
2.2.	Weighted Residual Approach.....	5
2.3.	Galerkin's Approach with One-Dimensional Hat Function	6
3.	Implementation of FEM under BS Model.....	9
3.1.	BS Equation	9
3.2.	European Type: Problem Formulation	10
3.3.	European Type: Example	12
3.4.	American Type: Linear Complementarity Problem	14
3.5.	American Type: Example	18
4.	Introduction to Heston's Model.....	21
5.	Implementation of FEM under Heston's Model	23
5.1.	Partial Differential Equation	23
5.2.	Green's First Identity and Functional Space	23
5.3.	Weak Formulation.....	24
5.4.	Mesh and Basis	25
5.5.	Reference Triangle and Affine Transformation	27
5.6.	Stiffness and Mass Matrix.....	28
5.7.	Solution and Discussion.....	31
6.	Conclusion	36

1. Introduction

The problem of option pricing plays a very important role both in modern financial theory and in practice. In the year of 1973, the Black-Scholes-Merton model theory was published and has gradually become the most widely used mathematical model for option pricing problems, for the reason that not only does it give an analytical solution for pricing the vanilla options, but also it is a strong foundation for the model at a more refined or extensional form. However, this model is based on several strict assumptions, which make it difficult to agree with the reality in financial market. For instance, one of the assumptions is that the continuously compounded returns are conformed to a normal distribution with constant volatility, which cannot be achieved usually. In retrospect to the history, the worldwide financial crisis in different time period show that return distributions exhibit significant skewness and kurtosis and have a negative correlation with implied volatility, instead of the normality assumption.

In the year of 1993, in order to extend constant volatility assumption to a more general and feasible one, an extension of Black-Scholes model was produced by Heston, in which the volatility was modeled by a stochastic process instead of being regarded as a constant. The Heston's model introduces a more accurate evaluation of financial derivatives. Also, for European options, it allows for the derivation of a closed-form exact solution whereas other stochastic volatility models can only be treated numerically.

It is usually hard to obtain closed-form solutions for some complicated partial differential equations (PDEs). Therefore, most option value computations are done by a numerical approximation.

The most commonly used method for solving PDE is the finite difference method, which requires some strict limitations such as sufficiently smooth terminal and boundary conditions, a rectilinear domain, and logically rectangular grids. The finite difference method with equidistant grids is easy to understand and straightforward to implement. It is easy to obtain the uniform rectangular grids, but not so flexible enough when applied to many situations. Steep gradients of the solution require for a finer grid so that the difference quotients can make a good approximations of the differentials. It is difficult to arrange such a flexible grid with finite difference method though it is possible to do that. In addition, the produced results are available only at the grid points.

In contrast, the finite element method, an alternative type of methods for solving the partial differential equations, can deal with a large variety of terminal and boundary conditions since it can provide high flexibility. The flexibility of this method is not only favorable to estimate functions, but also to approximate domains of computation that are not rectangular. Moreover, approximate solutions obtained are piecewise-smooth functions defined over all points in the solution domain. In this method, there are

several principles such as variational methods, weighted residuals, or Galerkin methods, which can serve to derive suitable equations.

The finite element method is of great importance to deal with the problems of multifactor options. For example, the Heston's model, a two dimensional problem concerning the influence of the time varying volatility, can be well approximated by the FEM. As for one-dimensional standard options, there are analytical forms of solutions which can be compared with the results from FEM so that it is worthy to take a deeper analysis on the improvement of this numerical method.

In this report, we mainly discuss the problem about the FEM application for option pricing with one-dimensional and two-dimensional PDEs to illustrate that the finite element method is more useful and efficiency when dealing with high-dimensional option pricing problems. Basically, we simply applied FEM to one-asset European option, and then made the boundary condition free to solve obstacle problem of the American option. Moreover, we broke the constant volatility assumption under the Black-Scholes model, pricing the European option under Heston's model and taking a comparison with its analytical solution. More detailed project problem will be further discussed in the following chapters.

2. Theory of Finite Element Method

As a preparation, in this section we sketch some important aspects of FEM without going ‘too deep’ mathematically. In section 2.1 we explain variational formulation approach first and then move to more practical approach called weighted residuals. Further, among choices of weighted residuals, we highlight Galerkin’s method as our focused method through this whole report.

To illustrate the spirit of FEM, we consider the following one dimensional problem with homogeneous Dirichlet condition for a domain $\Omega \in \mathbb{R}^n$ with boundary Γ :

$$\mathcal{P}_0: \begin{aligned} -\Delta u(\mathbf{x}) &= f(\mathbf{x}) && \text{in } \Omega \\ u(\mathbf{x}) &= 0 && \text{on } \Gamma \end{aligned} \quad (2.1)$$

where Δ is Laplace operator,

$$\Delta = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2} \quad (2.2)$$

2.1. Variational Formulation

Provided that $f(\mathbf{x}) \in \mathcal{C}^0(\bar{\Omega})$, a classical solution (or strong solution) of problem \mathcal{P}_0 is defined as a function $u \in \mathcal{C}^2(\bar{\Omega})$ satisfying relations specified in \mathcal{P}_0 . The theory of PDE tells us that more general counterpart problem of (2.1) will have unique solution provided all coefficient functions and f are sufficiently smooth.

Problem \mathcal{P}_0 can be reformulated so as to look for the solution in the distributional sense by testing the equation against smooth functions. Reformulating of the problem amounts to relaxing the pointwise continuity required to ensure the existence of the classical derivative to the existence of weaker derivative which continuity is to be interpreted in terms of Lebesgue spaces. The obtained problem is a weak formulation and a solution to this weak problem is thus called weak solution.

The test procedure of the problem condition (2.1) against any smooth function $\psi \in \mathcal{C}^\infty(\Omega)$ is given by:

$$-\int_{\Omega} \Delta u(\mathbf{x}) \psi(\mathbf{x}) d\mathbf{x} = \int_{\Omega} f(\mathbf{x}) \psi(\mathbf{x}) d\mathbf{x} \quad (2.3)$$

Since $u \in \mathcal{C}^2(\bar{\Omega})$, Δu is well defined. Integrating by parts and introducing gradient operator, the LHS reads:

$$-\int_{\Omega} \Delta u(\mathbf{x}) \psi(\mathbf{x}) d\mathbf{x} = \int_{\partial\Omega} \nabla u(\mathbf{x}) \cdot \mathbf{n} \psi(\mathbf{x}) ds + \int_{\Omega} \nabla u(\mathbf{x}) \cdot \nabla \psi(\mathbf{x}) d\mathbf{x} \quad (2.4)$$

The first line integral term vanishes because ψ has compact support in Ω . Thus the weak formulation reads:

$$\int_{\Omega} \nabla u(\mathbf{x}) \cdot \nabla \psi(\mathbf{x}) d\mathbf{x} = \int_{\Omega} f(\mathbf{x}) \psi(\mathbf{x}) d\mathbf{x}, \forall \psi \in \mathcal{C}^{\infty}(\Omega) \quad (2.5)$$

Then the weak formulation of problem \mathcal{P}_0 consists:

$$\mathcal{P}'_0: \text{Given } f \in V', \text{Find } u \in H, \text{s.t. } \int_{\Omega} \nabla u \cdot \nabla v d\mathbf{x} = \int_{\Omega} f v d\mathbf{x}, \forall v \in V \quad (2.6)$$

where H and V are functional spaces yet to be defined, satisfying continuity constraints and boundary condition constraints for H .

H and V are referred as solution space and test space, respectively. For test space V , since u vanishes on the boundary, if we choose $\psi \in \mathcal{H}_0^1(\Omega)$, then by definition, one can construct a sequence of functions $(\psi^n)_{n \in \mathbb{N}}$ in \mathcal{C}^{∞} converging in \mathcal{H}_0^1 to ψ . Consequently, the weak formulation is satisfied if $\psi \in \mathcal{H}_0^1(\bar{\Omega})$. As for the solution space, classical solution implies that $u \in \mathcal{C}^2(\bar{\Omega})$, which entails that $u \in L^2(\Omega)$ and $\partial_i u \in L^2(\Omega)$, thus $u \in \mathcal{H}^1(\Omega)$. Further, we can conclude that the solution $u \in \mathcal{H}_0^1(\Omega)$ with the help of Trace Theorem.

Another thing needs to be mentioned is that all derivatives involved in (2.6) should be understood as weak derivatives.

To fit in more general case, we introduce more abstract notation as following:

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v d\mathbf{x} \quad (2.7)$$

$$l(v) = \int_{\Omega} f v d\mathbf{x} \quad (2.8)$$

where $a(\cdot, \cdot)$ is a continuous bilinear form on $V \times V$ and $l(\cdot)$ is a continuous linear form on V . Then the general formulation, namely abstract problem, reads in this case:

$$\mathcal{P}: \text{Find } u \in V \text{ s.t. } a(u, v) = l(v), \forall v \in V \quad (2.9)$$

It can be showed that $a(u, v)$ and $l(v)$ defined above satisfy conditions specified by Lax-Milgram Theorem, which then guarantees the existence and uniqueness of the weak solution even in some cases that the functions involved in original problem setting are not smooth enough to permit a strong solution.

Ritz's method is based on replacing the solution space V (which is infinite dimensional) by a finite dimensional subspace $V_n \subset V$, $\dim(V_n) = n$. (2.9) is the approximate weak

problem by Ritz's method:

$$\mathcal{P}^h: \text{Find } u_n \in V_n \quad V_n \subset V. \quad \text{s. t. } a(u_n, v_n) = l(v_n) \quad \forall v_n \in V_n \quad (2.9)$$

Provided that the bilinear form is symmetric, (2.10) specifies an equivalent approximate variational problem under minimization form. It can be proved that if u is indeed the weak solution to (2.9) and under the assumption that $a(\cdot, \cdot)$ is symmetric, u is the unique minimizer of $J(\cdot)$ (defined as below) over V_n .

$$\begin{aligned} \mathcal{M}^h: \text{Find } u_n \in V_n \quad V_n \subset V. \quad \text{s. t. } J(u_n) \leq J(v_n) \quad \forall v_n \in V_n \\ \text{with } J(v_n) = \frac{1}{2}a(v_n, v_n) - l(v_n) \end{aligned} \quad (2.10)$$

We then can construct the solution:

$$u_n = \sum_i u_i \varphi_i \quad (2.11)$$

where u_i is a set of real numbers and $\mathcal{B} = \{\varphi_i\}_{i=1}^n$ a basis of V_n .

It can be proved that the minimization problem is well-posed and $u_n \rightarrow u$ in V as $n \rightarrow \infty$. The proof of convergence makes use of Cea's Lemma which we do not provide in this report detailed description for simplicity.

Alternatively, for the minimization problem to make more sense, one can view J in (2.10) a functional, which is the target of the minimization. The coefficients of u_n can be acquired by suppressing:

$$\frac{\partial J}{\partial u_i} = 0 \quad (2.12)$$

2.2. Weighted Residual Approach

The most difficult part of variational formulation is to find an appropriate functional associated with the problem to solve, whereas weighted residual approach is applicable to problems that the variational formulation is hard or impossible to achieve and thus offers more flexibility. In weighted residual approach, the idea is to construct a residual function based on the problem specified and then formulate the approximated problem by suppressing the weighted residual to be zero in an integral sense.

To illustrate, we still consider \mathcal{P}_0 posed in (2.1) and we follow the constructed solution which is represented by basis form as in (2.11). To repeat, we set approximate solution as $u_n = \sum u_i \varphi_i$. To determine the unknown coefficients, we need a residual function to apply weighted residual approach, which is defined as:

$$R = \Delta u_n - f \quad (2.13)$$

Since the basis function in space V_n are considered to be predefined, the residual R is

weighted by introducing n weighting functions $(\psi_i)_{i=1}^n$, which are also referred as testing functions as before. Then we seek solution by requiring:

$$\int_{\Omega} R\psi_j d\Omega = 0 \quad \forall 1 \leq j \leq n \quad (2.14)$$

which in fact formulate a system of equations equivalent to

$$\int_{\Omega} \Delta u_n \psi_j d\Omega = \int_{\Omega} f \psi_j d\Omega \quad \forall 1 \leq j \leq n \quad (2.15)$$

for the n unknown coefficients. The above equations can also be written using inner product notations as:

$$(\Delta u_n, \psi_j) = (f, \psi_j) \quad \forall 1 \leq j \leq n \quad (2.16)$$

Several different choices of weighting functions exist. A very popular choice and is used in this report is Bubnov-Galerkin method which defines $\psi_i = \varphi_i$, which means choosing weighting functions in the approximated weak solution space.

Another choice of weighting function worth to mention is Dirac's delta function. In \mathbb{R}^1 delta function satisfies:

$$\int f \delta(x - x_j) dx = f(x_j) \quad (2.17)$$

Consequently,

$$\int \Delta u_n \psi_j = \int \Delta u_n \delta_j = \int \Delta u_n x_j \quad (2.18)$$

$$\int f \delta_j = f(x_j) \quad (2.19)$$

This means that a system of equations $\Delta u_n(x_j) = f(x_j)$ results, which is equivalent to evaluating the specified PDE at selected points x_j . This method is often referred as collocation.

2.3. Galerkin's Approach with One-Dimensional Hat Function

Although FEM is often used to deal with high dimensional problem but we describe the basic procedures to implement Galerkin's method with the example of one-dimensional hat function as basis function. We now proceed under the assumption that $\bar{\Omega} = [0,1]$, for simplicity. The relevant domain is then subdivided into N subintervals $[x_{i-1}, x_i]$, $i = 1, \dots, N$, by the points $x_i = ih$, $i = 0, \dots, N$.

The prototype of a finite-element method makes use of the hat functions, which we define formally for $1 < i < m - 1$ as:

$$\varphi_i(x) = \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}} & \text{for } x_{i-1} < x < x_i \\ \frac{x - x_{i-1}}{x_i - x_{i-1}} & \text{for } x_i < x < x_{i+1} \\ 0 & \text{elsewhere} \end{cases} \quad (2.20)$$

To illustrate, we give a figure of one-dimensional hat functions below:

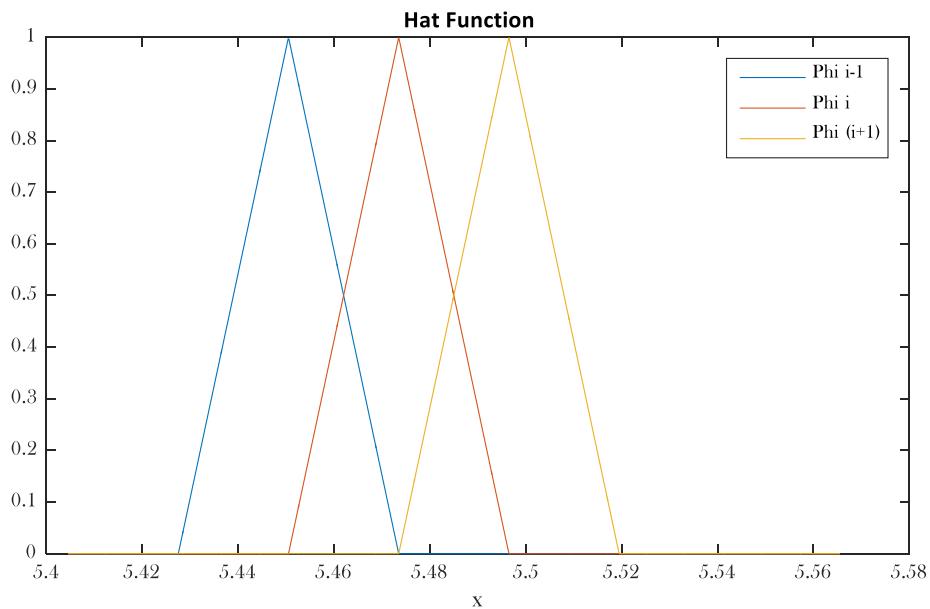


FIGURE 2.1 One-Dimensional Hat Functions

In our special case, boundary conditions suppress $u(0) = u(1) = 0$. But usually, special boundary functions are required to control for the boundary conditions. In one dimensional problem, boundary functions can be defined as:

$$\varphi_0(x) = \begin{cases} \frac{x_1 - x}{x_1 - x_{\min}} & \text{for } x_{\min} < x < x_1 \\ 0 & \text{elsewhere} \end{cases}$$

$$\varphi_m(x) = \begin{cases} \frac{x - x_{m-1}}{x_m - x_{m-1}} & \text{for } x_{m-1} < x < x_{\max} \\ 0 & \text{elsewhere} \end{cases} \quad (2.21)$$

Following (2.15), and using a finite dimensional space, i.e. $u_n = \sum u_i \varphi_i$, Galerkin's method results in a system of equations:

$$\sum_i u_i \int_{\Omega} \Delta \varphi_i \varphi_j dx = \int_{\Omega} f \varphi_j dx \quad \forall 0 \leq j \leq N \quad (2.22)$$

The integration on RHS will be simply approximated by substituting f by interpolating polygon:

$$f_p = \sum_i f(x_i) \varphi_i \quad (2.23)$$

Then RHS of (2.22) reads:

$$\int_{\Omega} f \varphi_j dx = \int_{\Omega} f_p \varphi_j dx = \sum_i f(x_i) \int_{\Omega} \varphi_i \varphi_j dx, \quad \forall j \quad (2.24)$$

Define $\int_{\Omega} \varphi_i \varphi_j dx = B_{ij}$ and $\int_{\Omega} \Delta \varphi_i \varphi_j dx = A_{ij}$, (2.22) becomes:

$$Au = B\bar{f} \quad (2.25)$$

where \bar{f} is a vector of $f(x_i)$.

In one-dimensional case, the Laplace operator reduces to ∂_{xx} . Integrating by parts leads to:

$$\int_{\Omega} \varphi_i'' \varphi_j dx = \int_{\Omega} \varphi_i' \varphi_j' dx \quad (2.26)$$

It's then trivial to get the mass matrix and stiffness matrix as (2.25). But it needs to be mentioned that A and B may be subjected to changes in order to incorporate boundary conditions since u_i can be seen as given at boundary nodes. This is especially that case in multidimensional case.

$$A = \frac{1}{h} \begin{bmatrix} 1 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{bmatrix} \quad B = \frac{h^2}{6} \begin{bmatrix} 2 & 1 & & & \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \\ & & & 1 & 2 \end{bmatrix} \quad (2.25)$$

3. Implementation of FEM under BS Model

Valuation of single asset European option with vanilla payoff structure makes use of Black-Scholes formula. But in order to test the efficiency and accuracy of FEM, following previous sections, we in this section apply FEM to classic BS model of option pricing. Section 3 is arranged as following: In 3.1 we summarize some well-known aspects of BS model to set the problem.

3.1. BS Equation

Black-Scholes (BS) model assumes that stock price follows a geometric Brownian motion. The dynamic under risk-neutral measure is given by:

$$dS_t = S_t(rdt + \sigma dW_t) \quad (3.1)$$

where r and σ are constant. We focus on the case where the stock is assumed to be non-dividend-paying but this assumption can be relaxed by including a constant continuous dividend rate in drift term, as what we do in section 5.

By Ito lemma, one can deduce the following well-known Black-Scholes equation for European option price $V(S, t)$:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0 \quad (3.2)$$

For constant parameters we can transform the equation to standard heat conduction equation by conducting the following change of variables:

$$S = Ke^x, \quad t = T - \frac{2\tau}{\sigma^2}, \quad q = \frac{2r}{\sigma^2} \quad (3.3)$$

$$\text{then } V(S, t) = v(x, \tau)$$

$$v(x, \tau) = K \exp \left\{ -\left(\frac{1}{2}(q-1)x\right) - \left(\frac{1}{4}(q+1)^2\tau\right) \right\} y(x, \tau) \quad (3.4)$$

Then solving the pricing PDE under BS model is equivalent to the following initial value problem but with domain $x \in (-\infty, +\infty)$ and $\tau \in [0, +\infty)$.

$$\frac{\partial y}{\partial \tau} = \frac{\partial^2 y}{\partial x^2} \quad (3.5)$$

with initial condition:

$$y_c(x, 0) = \gamma_c(x) = \max \left\{ \exp \left(\frac{x}{2} (q + 1) \right) - \exp \left(\frac{x}{2} (q - 1) \right), 0 \right\} \quad (3.6)$$

$$y_p(x, 0) = \gamma_p(x) = \max \left\{ \exp \left(\frac{x}{2} (q - 1) \right) - \exp \left(\frac{x}{2} (q + 1) \right), 0 \right\} \quad (3.7)$$

where subscripts c and p denote call and put, respectively.

We then truncate the x -domain and associated boundary conditions:

$$y_c(x_{min}, \tau) = \alpha_c(\tau) = 0 \quad (3.8)$$

$$y_c(x_{max}, \tau) = \beta_c(\tau) = \exp \left\{ \frac{1}{2}(q + 1)x_{max} + \frac{1}{4}(q + 1)^2\tau \right\} \quad (3.9)$$

$$y_p(x_{max}, \tau) = \beta_p(\tau) = 0 \quad (3.10)$$

$$y_p(x_{min}, \tau) = \alpha_p(\tau) = \exp \left\{ \frac{1}{2}(q - 1)x_{min} + \frac{1}{4}(q - 1)^2\tau \right\} \quad (3.11)$$

3.2. European Type: Problem Formulation

To make use of one-dimensional basis functions we apply $w = \sum w_i(\tau)\varphi_i(x)$ with the coefficient a function of τ . As a consequence of this simple approach, the same x -grid is applied to all τ , which results a rectangular grid in (x, τ) plane. Note that boundary condition $\beta(x, \tau)$ will reduces to a function just on τ after substitution of x_{max} in the truncated x -domian.

Taking into consideration of the boundary condition, the values w_0 and w_m would be known. We then introduce a special function to control the given Dirichlet boundary condition on function y . Then this special function no longer plays a role in solution space and is considered to be known. After this separation, the linear combination of basis does not reflect any none zero Dirichlet condition. The final ansatz becomes:

$$w = \sum_i w_i(\tau)\varphi_i(x) + \varphi_b(x, \tau) \quad (3.12)$$

where

$$\varphi_b(x, \tau) = (\beta(\tau) - \alpha(\tau)) \left(\frac{x - x_{min}}{x_{max} - x_{min}} \right) + \alpha(\tau) \quad (3.13)$$

We then choose basis functions to be hat functions defined earlier. One point needs to be mention is that because the introduction of the special basis $\varphi_b(x, \tau)$, the dimension of the unknown vector will be reduced by two compared with previous general setting.

It is trivial to calculate derivatives of the approximating function w and according to the problem specified in (3.5), we have the following equation by Galerkin's approach:

$$\int_{x_{\min}}^{x_{\max}} \left[\sum_{i=1}^{m-1} \frac{\partial w_i}{\partial \tau} \varphi_i + \frac{\partial \varphi_b}{\partial \tau} \right] \varphi_j dx = \int_{x_{\min}}^{x_{\max}} \left[\sum_{i=1}^{m-1} w_i \frac{\partial^2 \varphi_i}{\partial x^2} + \frac{\partial^2 \varphi_b}{\partial x^2} \right] \varphi_j dx \quad (3.14)$$

Arranging the terms and we can get matrix representation

$$Bw_\tau + b = -Aw - a \quad (3.15)$$

where B and A denotes mass matrix and stiffness matrix, respectively. Note that as mentioned before, we need to modify B and A a little bit because again the introduction of φ_b leads to unusual definition of the basis by forcing all other basis to zero at the boundary.

After taking in to the boundary, in this case we set up the following:

$$A = \frac{1}{h} \begin{bmatrix} 2 & -1 & & & 0 \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 2 & -1 \\ 0 & & & -1 & 2 \end{bmatrix} \quad B = \frac{h}{6} \begin{bmatrix} 4 & 1 & & & 0 \\ 1 & 4 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 4 & 1 \\ 0 & & & 1 & 4 \end{bmatrix} \quad (3.16)$$

$$b(\tau) = \begin{bmatrix} \int \frac{\partial \varphi_b}{\partial \tau} \varphi_1(x) dx \\ \vdots \\ \int \frac{\partial \varphi_b}{\partial \tau} \varphi_{m-1}(x) dx \end{bmatrix} \quad a(\tau) = \begin{bmatrix} \int \frac{\partial^2 \varphi_b}{\partial x^2} \varphi_1(x) dx \\ \vdots \\ \int \frac{\partial^2 \varphi_b}{\partial x^2} \varphi_{m-1}(x) dx \end{bmatrix} \quad (3.17)$$

For this specific φ_b we have $a = 0$ because $\partial_{xx} \varphi_b = 0$. For plain vanilla options the initial and boundary conditions are well defined and are given in (3.6-11). After setting up vector b analytically, (3.15) actually specifies a system of ordinary differential equations and is a method of lines since the solution along some level of x -domain is approximated by $w_i(\tau)$ as a function of τ .

By (3.6) and (3.7) we can enforce the initial condition by solving:

$$\sum_i w_i(0) \varphi_i(x) + \varphi_b(x, 0) = \gamma(x) \quad (3.18)$$

Since for any $x = x_i$, i.e. on the nodes, we have $\varphi_i(x) = 1$, the sum then reduces to

$$w_i(0) = \gamma(x_i) - \varphi_b(x_i, 0) \quad (3.19)$$

which serves as a value allocation for initial condition.

Till now we still have w_τ in hand and to achieve full discretization, we discretize τ -domain by Crank-Nicolson scheme. Denote $w(\tau)$ the vector function at node v by $w^{(v)}$, after some simple algebra and taking into consideration that $a = 0$ we reach:

$$\left(B + \frac{\Delta\tau}{2} A \right) w^{(v+1)} = \left(B - \frac{\Delta\tau}{2} A \right) w^{(v)} - \frac{\Delta\tau}{2} (b^{(v)} + b^{(v+1)}) \quad (3.20)$$

It can be readily seen that the above structure resembles FDM. In fact the convergence order is the same.

3.3. European Type: Example

In this section we examine a concrete example on call and put pricing and compare the solution with closed form solution given by BS formula. All the computation is done using MATLAB.

We set the option and computation parameters as follows:

TABLE 3.1: Parameters

Option Parameters		Computation Parameters	
Strike	$K = 100$	Spatial Resolution	$N = 1000$
Risk-free Rate	$r = 0.05$	Time Resolution	$M = 500$
Volatility	$\sigma = 0.4$		$x_{\min} = -13$
Time to Maturity	$T = 0.5$		$x_{\max} = 10$

We use the methods described above and get the following solution surface and important lines:

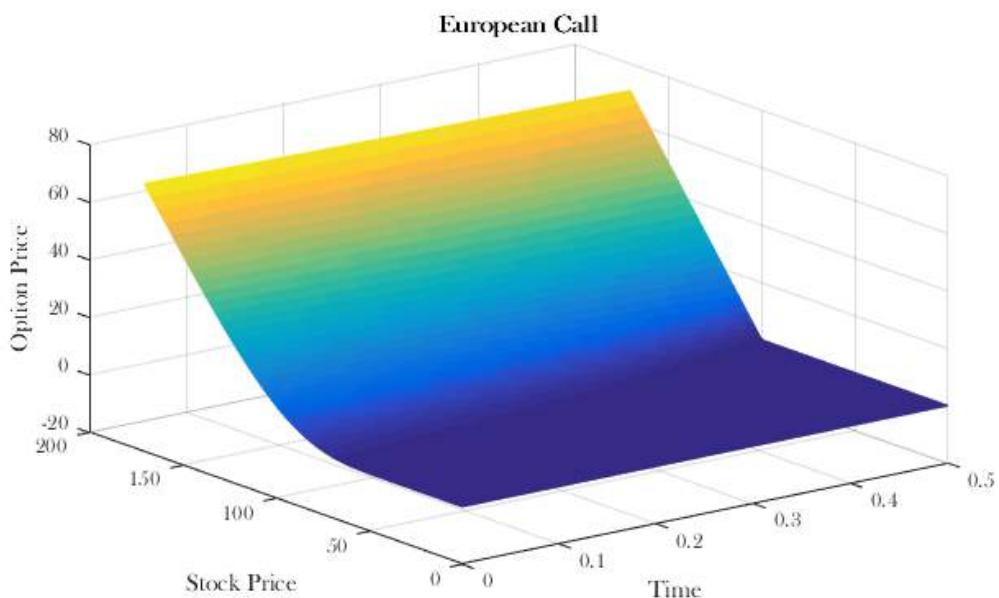


FIGURE 3.1 Solution Surface for European Call

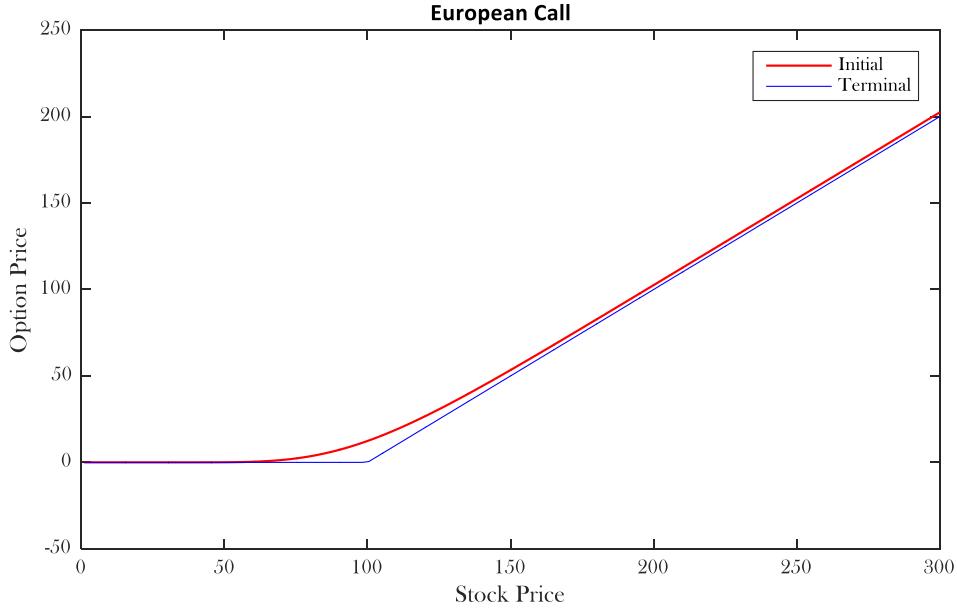


FIGURE 3.2: Comparison between Solutions at $T = 0$ and at $T = 0.5$ for European Call

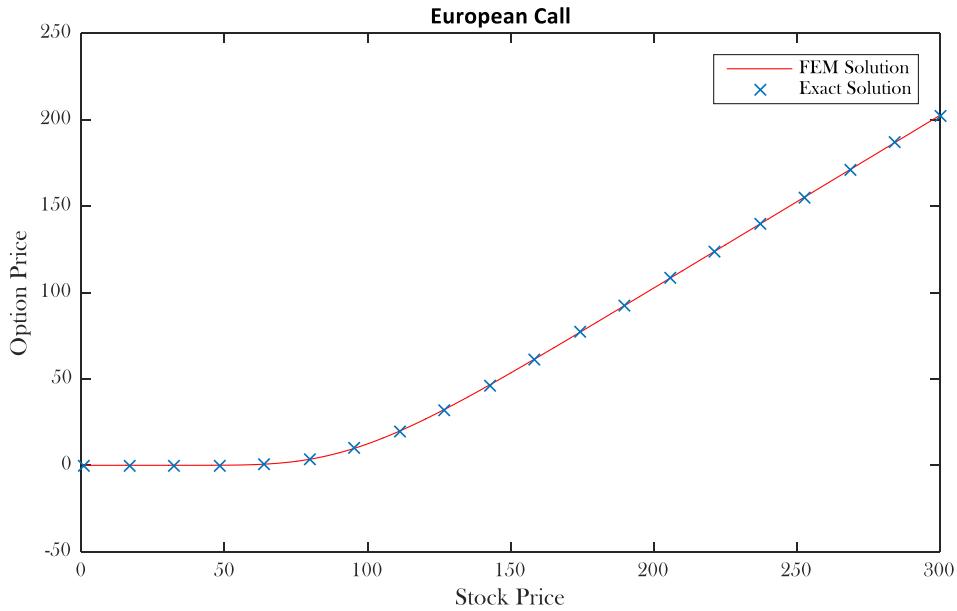


FIGURE 3.3: Comparison between FEM Solution and BS Solution for European Call

To better compare FEM solution with closed form BS solution, from the above picture we first take out some important points among the nodes of x -domain and compute their numerical values at $T = 0$.

TABLE 3.2: Numerical Comparison at the Nodes ($T = 0$)

Stock Price	FEM Solution	BS Solution	Relative Error
42.9491	0.0095921640	0.0098416658	-0.025351576
54.0433	0.1354307721	0.1364579953	-0.007527760
68.0033	1.0970402216	1.0982246862	-0.001078527
85.5693	5.3899836312	5.3884669649	0.000281465
107.6727	17.305881814	17.303084261	0.000161679
135.4858	39.911993542	39.912087798	-2.3616016e-06
170.4832	73.279695414	73.281115966	-1.9384982e-05
214.5209	117.02142196	117.02212723	-6.0268072e-06
269.9339	172.40442011	172.40471102	-1.6873639e-06

Then for arbitrarily tested points, we use linear interpolation to get approximated values, which are theoretically the same because the use of hat function leading to piecewise linear approximation between nodes. Numerical comparisons are presented as below:

TABLE 3.3: Numerical Comparison between FEM Solution and BS Solution ($T = 0$)

Stock Price	FEM Solution	BS Solution	Relative Error
80	3.5467894541	3.5463175338	0.00013307332
85	5.1796133065	5.1780812490	0.00029587360
90	7.2013850825	7.1993281385	0.00028571332
95	9.6099899724	9.6072338401	0.00028688090
100	12.388136207	12.385029207	0.00025086740
105	15.508750390	15.505722618	0.00019526798
110	18.938616847	18.935888150	0.00014410190
115	22.641344911	22.639024894	0.00010247864
120	26.580002293	26.578238481	6.63630258e-05

3.4. American Type: Linear Complementarity Problem

Different from the European type option, American option has a free boundary problem. Because there will be an early exercise as long as the payoff at each time is larger than the price of European option, the price of American option has to be at least the value of its payoff.

This problem can be summarized as follows:

$$\begin{aligned} AV_c(S, t) &\geq (S - K)_+ \quad \text{for all } (S, t) \\ AV_p(S, t) &\geq (K - S)_+ \quad \text{for all } (S, t) \end{aligned} \tag{3.21}$$

As for the American call, the value will not change comparing with the European type because the lower bound of the American option is lower than that of the European option. However, the case is different for the American put, since when the stock price drops to a certain level, the value of the European option will intersect with the lower

bound of American option at point $S_f(t)$.

$$\begin{aligned} AV_p(S, t) &\geq (K - S)_+ \quad \text{for } S > S_f(t) \\ AV_p(S, t) &= K - S \quad \text{for } S \leq S_f(t) \end{aligned} \tag{3.22}$$

Therefore, we have to further discuss the obstacle problem of the American put.

To illustrate we assume an obstacle $g(x)$. $g(x) > 0$ for $\alpha < x < \beta$, $g \in C^2$, $g'' < 0$ and $g(-1) < 0$, $g(1) < 0$. Since it is clear that American options are complementary in an analogous way:

$$\text{If } V > \text{payoff}, \quad \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0 \tag{3.23}$$

$$\text{If } V = \text{payoff}, \quad \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV < 0 \tag{3.24}$$

We adopted the linear complementarity theory for the American put option without dividends. The transformation leads to

$$\frac{\partial y}{\partial \tau} = \frac{\partial^2 y}{\partial x^2} \quad \text{as long as } AV_p(S, t) > (K - S)_+ \tag{3.25}$$

The relation of

$$AV_p(S, t) \geq (K - S)_+ = K \max\{1 - e^x, 0\} \tag{3.26}$$

leads to the following inequality:

$$\begin{aligned} y(x, \tau) &\geq \exp\left\{\frac{1}{2}(q-1)x + \frac{1}{4}(q+1)^2\tau\right\} \max\{1 - e^x, 0\} \\ &= \exp\left\{\frac{1}{4}(q+1)^2\tau\right\} \max\left\{(1 - e^x)\exp\left(\frac{1}{2}(q-1)x\right), 0\right\} \\ &= \exp\left\{\frac{1}{4}(q+1)^2\tau\right\} \max\left\{\exp\left(\frac{1}{2}(q-1)x\right) - \exp\left(\frac{1}{2}(q+1)x\right), 0\right\} \\ &= g(x, \tau) \end{aligned} \tag{3.27}$$

The function g is the obstacle function, which should satisfies the initial condition:

$$y(x, \tau) = g(x, \tau) \quad \text{for } x \rightarrow \infty \tag{3.28}$$

Therefore, based on the expression above, we can deal with the linear complementarity problem.

$$\begin{aligned} \left(\frac{\partial y}{\partial \tau} - \frac{\partial^2 y}{\partial x^2}\right)(y - g) &= 0 \\ \frac{\partial y}{\partial \tau} - \frac{\partial^2 y}{\partial x^2} &\geq 0, y - g \geq 0 \end{aligned} \tag{3.29}$$

$$\begin{aligned} y(x, 0) &= g(x, 0) \\ y(x_{\min}, \tau) &= g(x_{\min}, \tau), \quad y(x_{\max}, \tau) = g(x_{\max}, \tau) \end{aligned}$$

Before that, a function v out of a space \mathcal{K} is introduced:

$$\mathcal{K} = \{v \in C^0[-1, 1] : v(-1) = v(1) = 0, \quad v(x) \geq g(x) \text{ for } -1 \leq x \leq 1, \quad v \text{ piecewise } \in C^1\} \quad (3.30)$$

The function u can be compared to the function v so that the obstacle problem can be solved. It can be proved that if u solves the obstacle problem discussed above, then

$$\int_{-1}^1 u'(v - u)' dx \geq 0, \quad \forall v \in \mathcal{K} \quad (3.31)$$

which is related to a minimum problem, since when $v = u$ the integral will vanish.

As for the variational form of the American option, firstly the comparison functions must be redefined as:

$$\begin{aligned} \mathcal{K} = \{v \in C^0[x_{\min}, x_{\max}] : & \partial_x v \text{ piecewise } \in C^0, v(x, \tau) \geq g(x, \tau) \text{ for all } x, \tau, \\ & v(x, 0) = g(x, 0), \quad v(x_{\max}, \tau) = g(x_{\max}, \tau), \\ & v(x_{\min}, \tau) = g(x_{\min}, \tau)\} \end{aligned} \quad (3.32)$$

Then from the condition

$$v \geq g, \quad \frac{\partial y}{\partial \tau} - \frac{\partial^2 y}{\partial x^2} \geq 0 \quad (3.33)$$

we can deduce the final result of the variational form:

$$\int_{x_{\min}}^{x_{\max}} \left(\frac{\partial y}{\partial \tau} (v - y) + \frac{\partial y}{\partial x} \left(\frac{\partial v}{\partial x} - \frac{\partial y}{\partial x} \right) \right) dx \geq 0, \quad \forall v \in \mathcal{K} \quad (3.34)$$

Next is the implementation of pricing the American option with FEM method. In order to solve the minimum problem, we have to firstly take a discretization of y and v as:

$$y = \sum_i w_i(\tau) \varphi_i(x) \quad (3.35)$$

$$v = \sum_i v_i(\tau) \varphi_i(x) \quad (3.36)$$

With the finite element method, the x_i -grid is discretized by the basis function φ_i and for the time discretization, we adopted Crank-Nicolson method. Thus, after several steps of calculation, we can obtain the inequalities

$$(v^{(v+1)} - w^{(v+1)})^{tr} (B \frac{1}{\Delta\tau} (w^{(v+1)} - w^{(v)}) + \theta A w^{(v+1)} + (1 - \theta) A w^{(v)}) \geq 0 \quad (3.37)$$

The matrices A and B can be obtained from the way in solving the European option. θ is a parameter and when it takes the value of 0.5, it represents the Crank-Nicolson method. v denotes the step at the time-axis.

Set C and r as expressions as follows:

$$r = (B - \frac{1}{\Delta\tau} (1 - \theta) A) w^{(v)} \quad (3.38)$$

$$C = B + \Delta\tau\theta A \quad (3.39)$$

The inequality can be written as:

$$(v^{(v+1)} - w^{(v+1)})^{tr} (C w^{(v+1)} - r) \geq 0 \quad (3.40)$$

which is a completely discretization of the variational problem.

Further, the main idea of solving American option by finite elements is to construct w so that for all $v \geq g$:

$$(v - w)^{tr} (C w - r) \geq 0, w \geq g \quad (3.41)$$

This condition can be transformed into the way of finite difference method by substituting A by C and b by r . And thus the inequality becomes:

$$C w - r \geq 0, w \geq g \quad (3.42)$$

$$(C w - r)^{tr} (w - g) = 0 \quad (3.43)$$

It can be proved that an equivalence hides between the problem in FEM and FDM.

The sufficient condition from FDM to FEM is true because in FDM the function w must satisfy $w \geq g$, so that the inequality in FEM can be written as follows

$$(v - w)^{tr} (C w - r) = (v - g)^{tr} (C w - r) - (w - g)^{tr} (C w - r) \geq 0 \quad (3.44)$$

in which the second term of the left-hand side equals to zero. Thus,

$$(v - g)^{tr} (C w - r) \geq 0 \text{ for all } v \geq g \quad (3.45)$$

The necessary condition is also true because:

$$v^{tr} (C w - r) \geq w^{tr} (C w - r) \text{ for all } v \geq g \quad (3.46)$$

and if there is a component of $C w - r$ is negative and make the corresponding step of v arbitrarily large. Then the LHS will be arbitrarily small, which is a contradiction. Therefore, it must be satisfied that $C w - r \geq 0$

And $w \geq g$ leads to

$$(w - g)^{tr}(Cw - r) \geq 0 \quad (3.47)$$

Let $v = g$ in FEM, then

$$(w - g)^{tr}(Cw - r) \leq 0 \quad (3.48)$$

Consequently, we can get

$$(w - g)^{tr}(Cw - r) = 0 \quad (3.49)$$

As a consequence of this equivalence, the process of solving American option under FEM is actually equivalent with that under FDM, which means the solutions in finite element method can be computed with the method in finite difference.

At each node, after the comparison, the discretized form $w(x, \tau)$ is substituted by $g(x, \tau)$ when

$$w(x, \tau) < g(x, \tau) \quad (3.50)$$

which is similar with the process of FDM. But the main difference is that $w(x, \tau)$ is numerically solved by FEM and when the substitution is done, the new value is used to calculate the next discretized value by the way of solving the European option.

3.5. American Type: Example

From the result, we can easily find that the price of the American option never drops below the lower bound $\max(K - S, 0)$. When $\tau = 0$, which means that there is no time value in the option, the option price will therefore get identical with the value of the lower bound with a kink at the point of the exercise price. If the time value is considered, the curve of the American put option will be smoother. Moreover, it's worthy to mention that when the stock price drops below $S_f(t)$, the curve overlaps the lower bound, which can prevent investors from exercising the option early and that's why the value of American put option is usually larger than that of European put option.

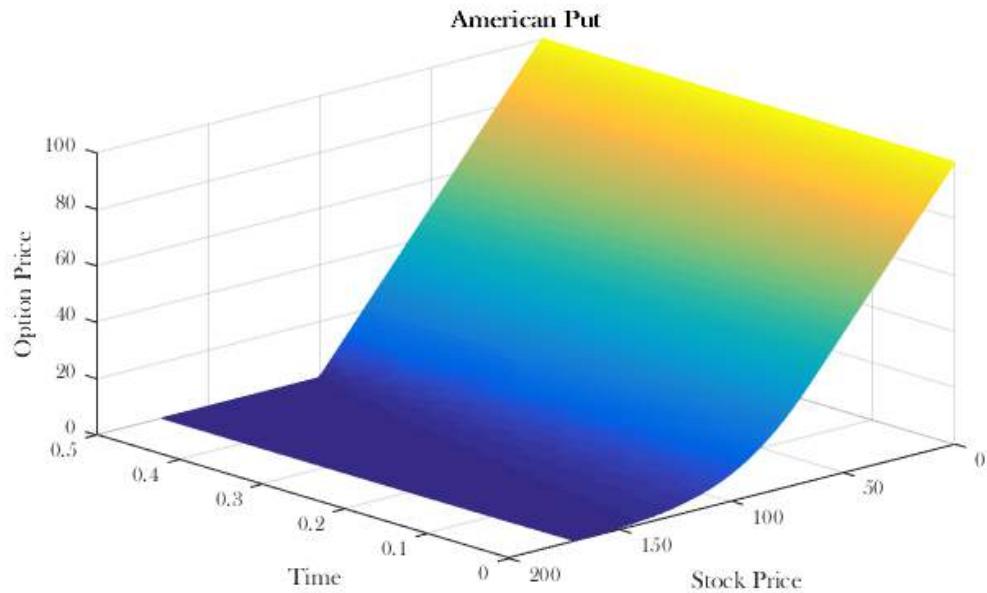
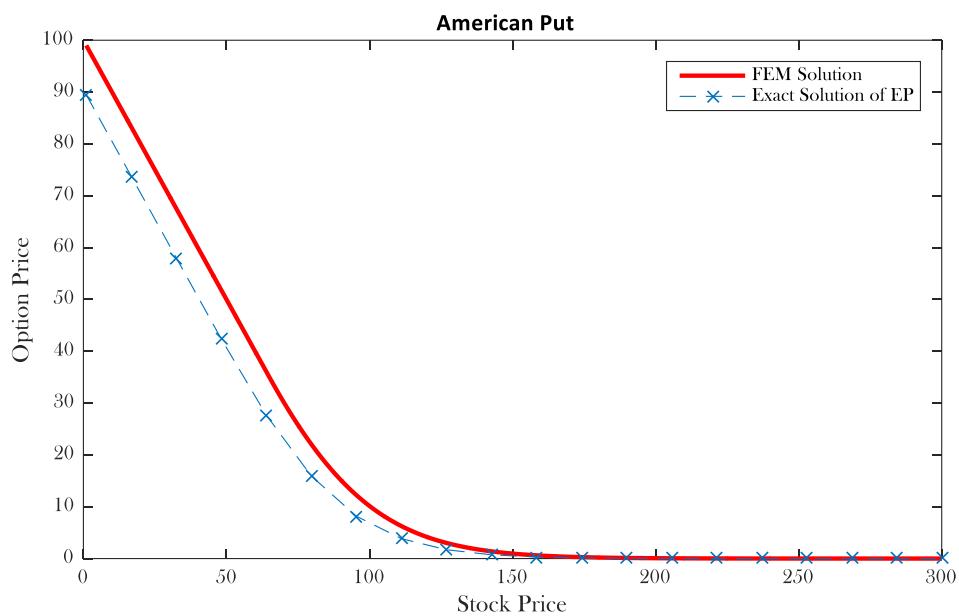


FIGURE 3.4: Solution Surface for American Put



**FIGURE 3.5: Comparison between BS Solution for European Put
and FEM Solution for American Put**

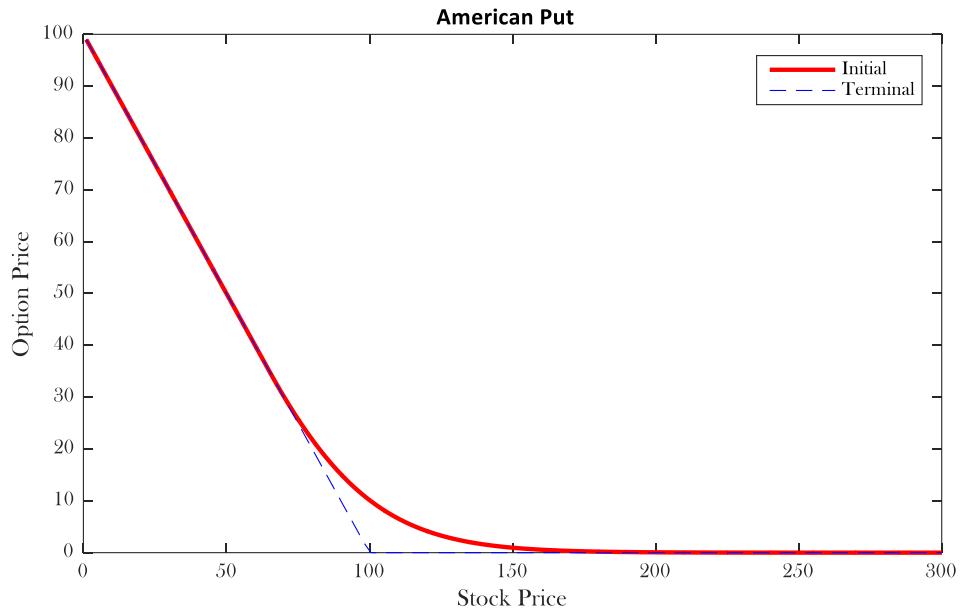


FIGURE 3.6: Comparison between Solutions at $T = 0$ and at $T = 0.5$ for American Call

4. Introduction to Heston's Model

Black-Scholes model has become the most widely used mathematical model for option pricing problem. Although BS model makes quite simple assumptions on price behavior of the stock, it does serve as a solid foundation for more sophisticated and extensional models. We then have a class of so-called stochastic volatility models, among which the most popular is Heston's model.

The Heston's model is formally defined as a system of SDEs given by:

$$\begin{aligned} dS_t &= S_t [(r - q)dt + \sqrt{v_t}dW_{1t}] \\ dv_t &= \kappa(\theta - v_t)dt + \xi\sqrt{v_t}dW_{2t} \\ dW_{1t}dW_{2t} &= \rho dt \end{aligned} \tag{4.1}$$

where the S_t denotes the spot process at time t and v_t the instantaneous volatility.

The parameters ρ , ξ and κ , which are included in Heston's model, provide the ability to capture observed features of the market and to produce a wide range of distributions of spot process. For example, ρ , the correlation between log-returns and the volatility, affects the skewness of the price distribution and hence have a great impact on the shape of implied volatility surface; volatility of the variance ξ , affects the kurtosis of the distribution; the mean reversion speed κ , can be seen as the degree of volatility clustering, which has been constantly observed in the market.

By Fundamental Theorem of Asset Pricing we can denote the option price with volatility v and spot price S as

$$g(t, v, S) = e^{-r(T-t)}\mathbb{E}[h(V_T, S_T)] \tag{4.2}$$

where $h(V_T, S_T)$ is the payoff of the option at maturity time T .

By the Feynman-Kac Theorem we then have that the function $g(t, v, S)$ satisfies the following PDE:

$$\frac{\partial g}{\partial t} + \frac{1}{2}\xi^2 v \frac{\partial^2 g}{\partial v^2} + \rho\xi S v \frac{\partial^2 g}{\partial S \partial v} + \frac{1}{2}S v \frac{\partial^2 U}{\partial S^2} + \kappa(\theta - v) \frac{\partial g}{\partial v} + (r - q)S \frac{\partial g}{\partial S} - rg = 0 \tag{4.3}$$

with the condition that $g(t, v, S) = h(v, S)$ imposed at the maturity time T .

The value $g(0, v, S)$ at the initial time are unknown and need to be determined. By variable conversion of $y = \log(S/K)$ and $\tau = T - t$, resulting in:

$$\frac{\partial U}{\partial \tau} - \frac{1}{2}\xi^2 v \frac{\partial^2 U}{\partial v^2} - \rho\xi v \frac{\partial^2 U}{\partial y \partial v} - \frac{1}{2}v \frac{\partial^2 U}{\partial y^2} - \kappa(\theta - v) \frac{\partial U}{\partial v} - (r - q - \frac{1}{2}v) \frac{\partial U}{\partial y} + rU = 0 \tag{4.4}$$

Simultaneously, terminal condition becomes initial condition in the time to maturity domain as $U(0, v, y) = h(v, Ke^y)$.

In order to clearly pose the pricing problem in PDE setting we need to specify boundary conditions. We constrain the computation domain in $\Omega = \{(v, y) : v \in (v_{\min}, v_{\max}), y \in (y_{\min}, y_{\max})\}$, where $y_{\min} = \log(S_{\min}/K)$ and $y_{\max} = \log(S_{\max}/K)$ with the following boundaries:

$$\begin{aligned}\Gamma_1 &= \{(v, y) : v = v_{\min}, y \in (y_{\min}, y_{\max})\} \\ \Gamma_2 &= \{(v, y) : y = y_{\max}, v \in (v_{\min}, v_{\max})\} \\ \Gamma_3 &= \{(v, y) : v = v_{\max}, y \in (y_{\min}, y_{\max})\} \\ \Gamma_4 &= \{(v, y) : y = y_{\min}, v \in (v_{\min}, v_{\max})\}\end{aligned}\tag{4.5}$$

Then the Dirichlet boundary condition can be specified as:

$$\begin{aligned}U(t, v_{\min}, y) &= [\eta(Ke^y e^{-q\tau} - Ke^{-r\tau})]^+ \\ U(t, v, y_{\max}) &= \frac{1+\eta}{2} [\eta(Ke^{y_{\max}} e^{-q\tau} - Ke^{-r\tau})]^+ \\ U(t, v_{\max}, y) &= \frac{1+\eta}{2} Ke^y e^{-q\tau} + \frac{1-\eta}{2} Ke^{-r\tau} \\ U(t, v, y_{\min}) &= \frac{1-\eta}{2} [\eta(Ke^{y_{\min}} e^{-q\tau} - Ke^{-r\tau})]^+ \\ U(0, v, y) &= [\eta(Ke^y - K)]^+\end{aligned}\tag{4.6}$$

To rewrite (4.3) we introduce gradient and divergence operators:

$$\partial_\tau U - \nabla \cdot A \nabla U + b \cdot \nabla U + rU = 0\tag{4.7}$$

where

$$A = \begin{bmatrix} \frac{1}{2}v\xi^2 & \alpha v \rho \xi \\ (1-\alpha)v\rho\xi & \frac{1}{2}v \end{bmatrix} \quad b = \begin{bmatrix} -\kappa(\theta - v) + \frac{1}{2}\xi^2 \\ -(r - q) + \frac{1}{2}v + \alpha \rho \xi \end{bmatrix}\tag{4.8}$$

Choosing $\alpha = 0.5$ leads to usual formulation of Heston's model presented in most literatures. For computational simplicity it may be beneficial to use $\alpha = 0$ or $\alpha = 1$.

5. Implementation of FEM under Heston's Model

5.1. Partial Differential Equation

We can note that gradient operator $\nabla = \begin{pmatrix} \partial_v \\ \partial_y \end{pmatrix}$ and (\cdot) denotes dot product. From previous (4.7) and (4.8) we can get

$$\nabla \cdot A \nabla = \begin{pmatrix} \partial_v \\ \partial_y \end{pmatrix}^{\text{tr}} \frac{1}{2} v \begin{bmatrix} \xi^2 & \rho\xi \\ \rho\xi & 1 \end{bmatrix} \begin{pmatrix} \partial_v \\ \partial_y \end{pmatrix} \quad (5.1.1)$$

$$\nabla \cdot A \nabla = \frac{\xi^2}{2} \partial_v + \frac{\rho\xi}{2} \partial_y + \rho\xi v \partial_{vy} + \frac{v\xi^2}{2} \partial_{vv} + \frac{v}{2} \partial_{yy} \quad (5.1.2)$$

$$b \cdot \nabla = \begin{pmatrix} \frac{\xi^2}{2} - \kappa(\theta - v) \\ \frac{v + \rho\xi}{2} - (r - q) \end{pmatrix}^{\text{tr}} \begin{pmatrix} \partial_v \\ \partial_y \end{pmatrix} \quad (5.2.1)$$

$$b \cdot \nabla = \left(\frac{\xi^2}{2} - \kappa(\theta - v) \right) \partial_v + \left(\frac{v + \rho\xi}{2} - (r - q) \right) \partial_y \quad (5.2.2)$$

The partial derivative respect to time is approximated by finite difference method, taking use of Crank-Nicolson scheme. Equation (5.1) can be expressed as

$$\frac{U^{k+1} - U^k}{d\tau} - \frac{1}{2} (\nabla \cdot A \nabla - b \cdot \nabla - r)(U^{k+1} + U^k) = 0 \quad (5.3)$$

where $d\tau = T/N_\tau$ and the partition of time is $\tau^0 = 0 < \tau^1 < \dots < \tau^{N_\tau} = T$. At each time step, U^k is computed from previous step and U^{k+1} can be solved from (5.3). After the rearrangement, we can get following equations

$$\begin{aligned} \mathcal{L}U^k &= \mathcal{R}U^{k+1} \\ \mathcal{L} &= \frac{1}{d\tau} + \frac{1}{2} (\nabla \cdot A \nabla - b \cdot \nabla - r) \\ \mathcal{R} &= \frac{1}{d\tau} - \frac{1}{2} (\nabla \cdot A \nabla - b \cdot \nabla - r) \end{aligned} \quad (5.4)$$

5.2. Green's First Identity and Functional Space

The first Green's identity is derived from the Divergence Theorem: consider the scalar function ϕ' and φ' on the domain $\Omega' \subset \mathbb{R}^d$, in which ϕ' is twice continuously differentiable and φ' is once continuously differentiable. Then

$$\int_{\Omega} (\varphi' \Delta \phi' + \nabla \varphi' \cdot \nabla \phi') dV = \oint_{\partial\Omega} \varphi' (\Delta \phi' \cdot \mathbf{n}) d\Gamma \quad (5.5)$$

where Γ denotes piecewise smooth boundary and \mathbf{n} denotes outward unit surface normal to Γ .

We also need to define the functional space in this problem for the solution U and test function ψ as follow

$$\begin{aligned} \mathcal{L}^2(\Omega) &:= \{f: \|f\|_{\mathcal{L}^2} < \infty\}, \text{ where } \|f\|_{\mathcal{L}^p} = \left(\int_{\Omega} |f|^p d\Omega \right)^{\frac{1}{p}} \\ \mathcal{H}^1(\Omega) &:= \{f \in \mathcal{L}^2(\Omega): \mathcal{D}^1 f \in \mathcal{L}^2(\Omega)\} \\ \mathcal{H}_0^1(\Omega) &= \{f \in \mathcal{H}^1(\Omega): f = 0 \text{ on } \Gamma\} \\ \mathcal{H}_b^1(\Omega) &= \{f \in \mathcal{H}^1(\Omega): f = g \text{ on } \Gamma\} \end{aligned} \quad (5.6)$$

$\mathcal{D}^1 f$ denotes the first order weak partial derivative of function f . $\mathcal{L}^2(\Omega)$ is the Lebesgue space with Euclidean norm which coincide with Hilbert Space in this case. The $\mathcal{H}^1(\Omega)$, $\mathcal{H}_0^1(\Omega)$, $\mathcal{H}_b^1(\Omega)$ are all Sobolev Space.

5.3. Weak Formulation

Equation (5.7) can be transferred into a weak formulation by multiplying on the both side with a scalar-valued test function ψ , and we can get the final pricing problem posed as:

$$\int_{\Omega} \psi \mathcal{L} U^k d\Omega = \int_{\Omega} \psi \mathcal{R} U^{k+1} d\Omega \quad (5.7.1)$$

$$\begin{aligned} \text{LHS} &= \frac{1}{2} \oint_{\Gamma} \psi A \nabla U^k \cdot \mathbf{n} d\Gamma - \frac{1}{2} \int_{\Omega} \nabla \psi \cdot A \nabla U^k d\Omega - \frac{1}{2} \int_{\Omega} \psi b \cdot \nabla U^k d\Omega \\ &\quad + \left(\frac{1}{d\tau} - \frac{r}{2} \right) \int_{\Omega} \psi U^k d\Omega \end{aligned} \quad (5.7.2)$$

$$\begin{aligned} \text{RHS} &= -\frac{1}{2} \oint_{\Gamma} \psi A \nabla U^{k+1} \cdot \mathbf{n} d\Gamma + \frac{1}{2} \int_{\Omega} \nabla \psi \cdot A \nabla U^{k+1} d\Omega \\ &\quad + \frac{1}{2} \int_{\Omega} \psi b \cdot \nabla U^{k+1} d\Omega + \left(\frac{1}{d\tau} + \frac{r}{2} \right) \int_{\Omega} \psi U^{k+1} d\Omega \end{aligned} \quad (5.7.3)$$

The variational form (5.7.1) is also called the weak form, in which the requirements of the solution have been considerably weakened over the strong form. In the weak formulation of the problem, the trial solution and the test function must belong to $\mathcal{H}^1(\Omega)$, which means that it is not necessary that all functions and derivatives be continuous.

We use the inhomogeneous Dirichlet boundary condition $U = g$ on the boundary Γ . And assume that there is a function $G \in \mathcal{H}^1(\Omega)$ so that $G = g$ on the boundary Γ and function $W \in \mathcal{H}_0^1(\Omega)$ so that solution $U = W + G$ and $U \in \mathcal{H}_b^1(\Omega)$.

The boundary integral appeared in the first term in (5.7.2) and (5.7.3) vanishes because $\psi \in \mathcal{H}_0^1$ is zero on the boundary Γ and we get:

$$\begin{aligned} & -\frac{1}{2} \int_{\Omega} \nabla \psi \cdot A \nabla U^k d\Omega - \frac{1}{2} \int_{\Omega} \psi b \cdot \nabla U^k d\Omega + \left(\frac{1}{d\tau} - \frac{r}{2} \right) \int_{\Omega} \psi U^k d\Omega = \\ & \frac{1}{2} \int_{\Omega} \nabla \psi \cdot A \nabla (W^{k+1} + G^{k+1}) d\Omega + \frac{1}{2} \int_{\Omega} \psi b \cdot \nabla (W^{k+1} + G^{k+1}) d\Omega \quad (5.8) \\ & + \left(\frac{1}{d\tau} + \frac{r}{2} \right) \int_{\Omega} \psi (W^{k+1} + G^{k+1}) d\Omega \end{aligned}$$

5.4. Mesh and Basis

We denote the close domain $\bar{\Omega} = \Omega + \Gamma = [v_{\min}, v_{\max}] \times [y_{\min}, y_{\max}]$, which is decomposed into mesh consisting of triangles T_i , $i = 1, \dots, N^T$ and their nodes D_i , $i = 1, \dots, N^D$. We assume that the mesh consists $N \times M$ squares and the total number of triangles is $N^T = 2MN$ and total number of nodes is $N^D = (N + 1)(M + 1)$, where N and M are spatial resolution of variable v and y . And we label the triangles from the bottom in the upward direction to the top in each column, which is from the left to the right. At the same time, we label the nodes from the left to the right in each row, which is from the bottom to the top.

A triangle basis (Lagrange basis) φ_i , $i = 1, \dots, N^D$ is used and we have $\varphi_i = 1$ at node D_i and $\varphi_i = 0$ at all other nodes, as presented below.

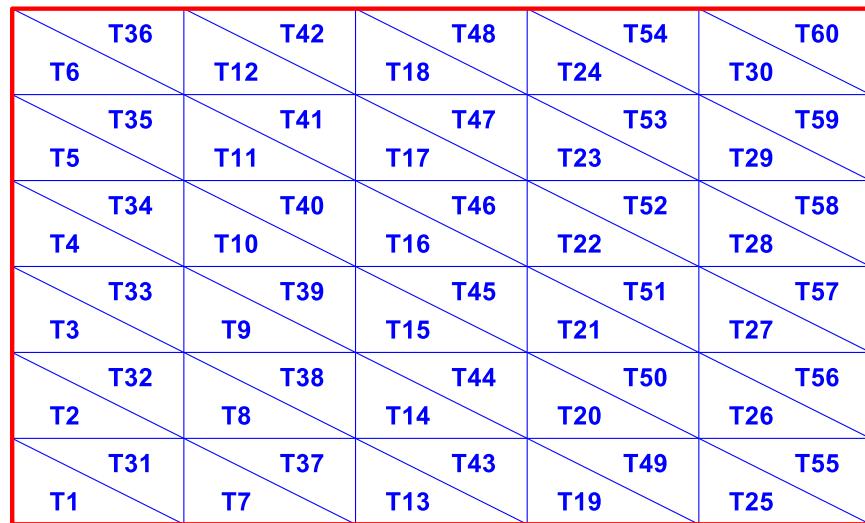


FIGURE 5.1: Label of the Triangle

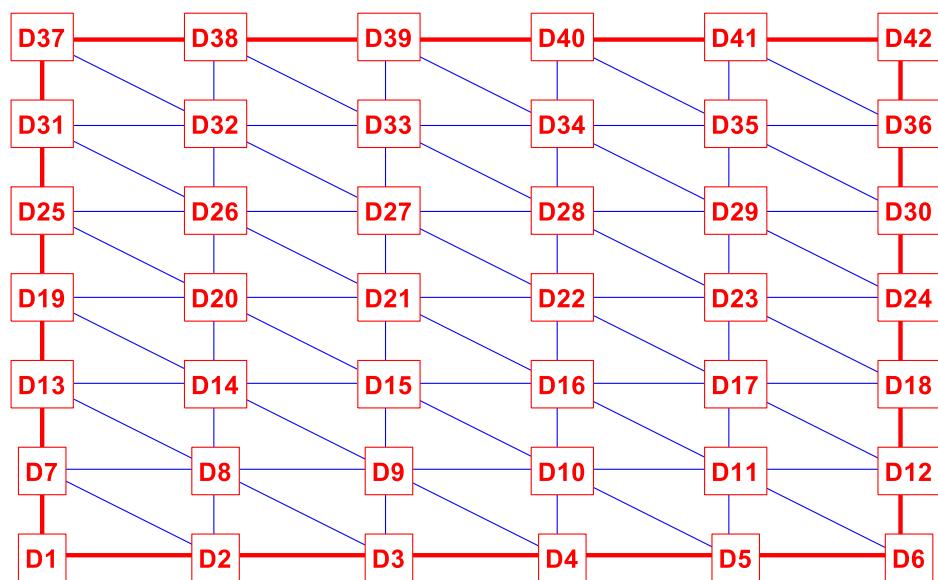
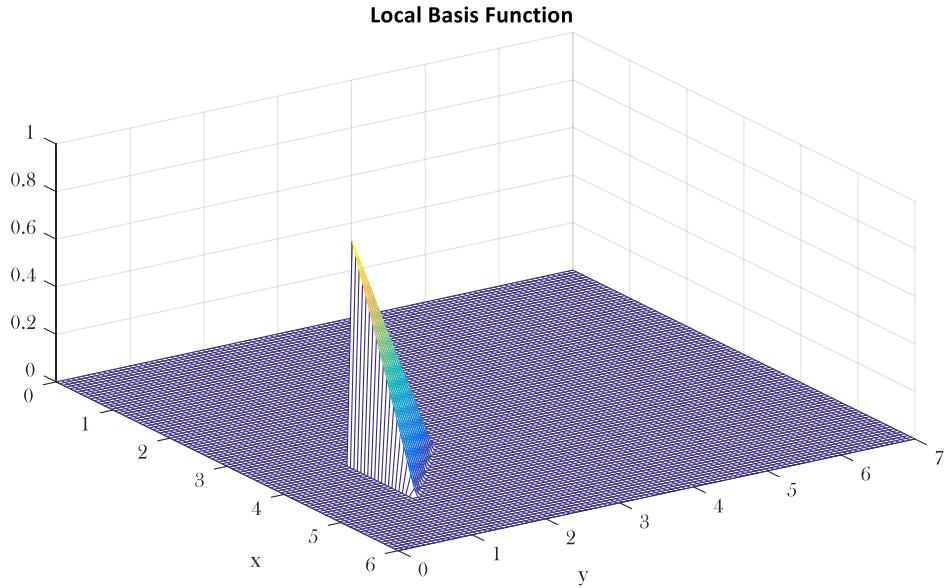
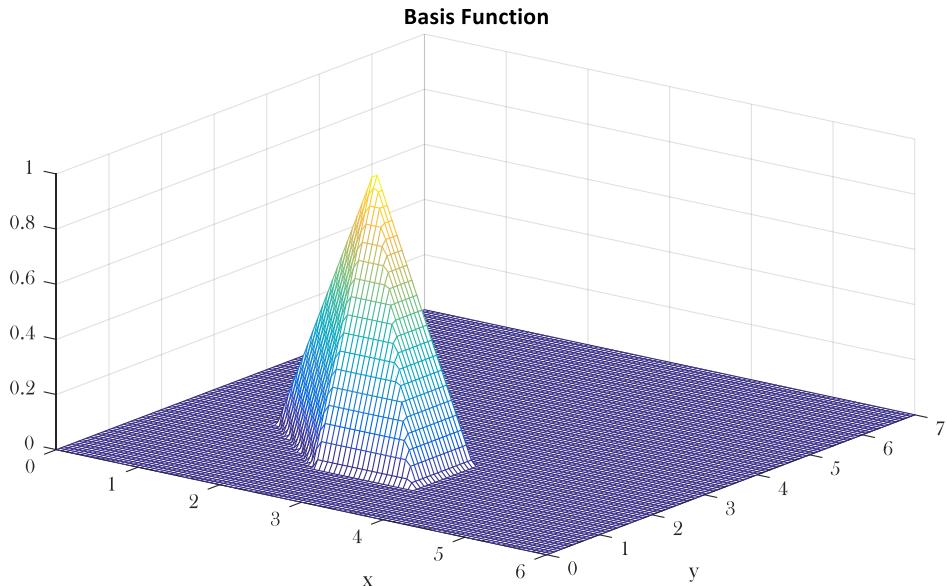


FIGURE 5.2: Label of the Nodes

**FIGURE 5.3: Local Basis Function****FIGURE 5.4: Element Function**

5.5. Reference Triangle and Affine Transformation

Consider a reference triangle $\hat{T} = \{(\hat{v}, \hat{y}) \in \mathbb{R}^2 : \hat{v} \in [0,1], \hat{y} \in [0,1 - \hat{v}]\}$, where a hat is used to denote the variable defined in reference triangle. The three nodes in the reference triangles D_i , $i = 0,1,2$ are $(0,0)$, $(1,0)$, $(0,1)$, respectively. In the figure, the open circle denotes $\hat{\varphi}_i = 0$ at nodes D_j if $i \neq j$ while the black dot denotes $\hat{\varphi}_i = 1$ at nodes D_j is $i = j$. Therefore, the three basis function on the reference

triangle are $\hat{\phi}_0(\hat{v}, \hat{y}) = 1 - \hat{v} - \hat{y}$, $\hat{\phi}_1(\hat{v}, \hat{y}) = \hat{v}$, $\hat{\phi}_2(\hat{v}, \hat{y}) = \hat{y}$. We then map the reference triangle to the arbitrage triangle T in our mesh with three nodes $D_0(v_0, y_0)$, $D_1(v_1, y_1)$, $D_2(v_2, y_2)$, resorting to the following affine transformation:

$$\begin{pmatrix} v \\ y \end{pmatrix} = F(\hat{v}, \hat{y}) = J \begin{pmatrix} \hat{v} \\ \hat{y} \end{pmatrix} + \begin{pmatrix} v_0 \\ y_0 \end{pmatrix}, \text{ where } J = \begin{bmatrix} v_1 - v_0 & v_2 - v_0 \\ y_1 - y_0 & y_2 - y_0 \end{bmatrix} \quad (5.9)$$

Again, the appearance of local basis function on this reference triangle is the same as presented in Figure 5.3. Below we give the corresponding local basis function defined on the reference triangle:

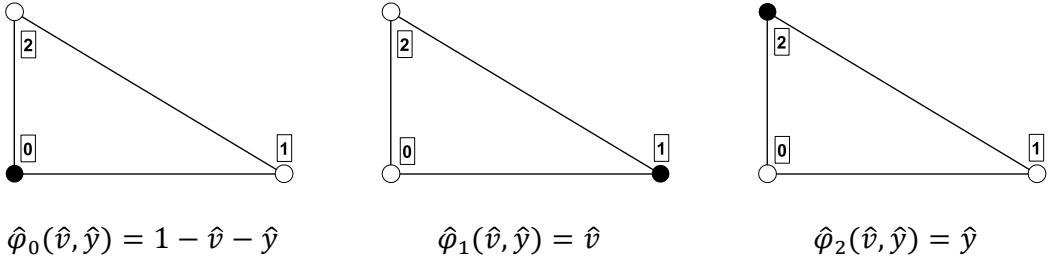


FIGURE 5.5: Three Basis Function on the Reference Triangle

By employing this affine transformation, actually we are treating any triangular domain in Ω as the image under F . This treatment is both to match the local indices of three vertices to global indices and to perform interpolation when bilinear interpolation is not appropriate. In more general cases, when the mesh is unstructured, the point coordinate can be converted to a barycentric coordinate in any triangular domain to get interpolated function value at an arbitrary position in the domain.

5.6. Stiffness and Mass Matrix

Our problem involves inhomogeneous Dirichlet boundary condition $U = g$ on the boundary Γ . We can approximate the target function by the continuous piecewise linear function G by using triangle basis φ_i , which is defined as

$$G = \begin{cases} 0 & \text{if } D_i \notin \Gamma \\ g_i \varphi_i & \text{if } D_i \in \Gamma \end{cases} \quad \text{where } g_i = g(D_i) \quad (5.10)$$

Thus, at time τ^k , the trial solution $U^k = W^k + G^k$, and

$$G^k = \sum_{D_i \in \Gamma} g_i^k \varphi_i, \quad W^k = \sum_{D_i \in \Omega} W_i^k \varphi_i \quad (5.11)$$

And the test function $\psi \in \mathcal{H}_0^1$ can be chosen as a linear combination of the basis function, which is

$$\psi = \sum_{D_i \in \Omega} b_i^k \varphi_i \quad (5.12)$$

where b_i is arbitrarily chosen. We can then simplify the test function as:

$$\psi = \sum_{D_i \in \Omega} \varphi_i \quad (5.13)$$

Therefore, similar to most general parabolic PDE, we get

$$\begin{aligned} & \int_{\Omega} \left(\sum_{D_i \in \Omega} \varphi_i \cdot \right) \mathcal{L} \left(\sum_{D_i \in \Gamma} g_i^k \varphi_i + \sum_{D_i \in \Omega} W_i^k \varphi_i \right) d\Omega \\ &= \int_{\Omega} \left(\sum_{D_i \in \Omega} \varphi_i \cdot \right) \mathcal{R} \left(\sum_{D_i \in \Gamma} g_i^{k+1} \varphi_i + \sum_{D_i \in \Omega} W_i^{k+1} \varphi_i \right) d\Omega \end{aligned} \quad (5.14)$$

which can be write as:

$$SW^k + PG^k = MW^{k+1} + NG^{k+1}$$

$$S_{ij} = \int_{\Omega} \varphi_i \mathcal{L} \varphi_j d\Omega \quad \forall D_i, D_j \in \Omega$$

$$P_{ij} = \int_{\Omega} \varphi_i \mathcal{L} \varphi_j d\Omega \quad \forall D_i \in \Omega, D_j \in \Gamma \quad (5.15)$$

$$M_{ij} = \int_{\Omega} \varphi_i \mathcal{R} \varphi_j d\Omega \quad \forall D_i, D_j \in \Omega$$

$$N_{ij} = \int_{\Omega} \varphi_i \mathcal{R} \varphi_j d\Omega \quad \forall D_i \in \Omega, D_j \in \Gamma$$

where matrices S , P , M , N are matrices of size $(f \times f)$, $(f \times c)$, $(f \times f)$, $(f \times c)$ respectively, where f and c denote the number of free nodes and constrained nodes.

The computation of the two integrals in the following is very crucial, which can be expressed as:

$$\begin{aligned} \int_{\Omega} \varphi_i \mathcal{L} \varphi_j d\Omega &= -\frac{1}{2} \int_{\Omega} \nabla \varphi_i \cdot A \nabla \varphi_j d\Omega - \frac{1}{2} \int_{\Omega} \varphi_i b \cdot \nabla \varphi_j d\Omega \\ &\quad + \left(\frac{1}{d\tau} - \frac{r}{2} \right) \int_{\Omega} \varphi_i \varphi_j d\Omega \end{aligned} \quad (5.16.1)$$

$$\begin{aligned} \int_{\Omega} \varphi_i \mathcal{R} \varphi_j d\Omega &= \frac{1}{2} \int_{\Omega} \nabla \varphi_i \cdot A \nabla \varphi_j d\Omega + \frac{1}{2} \int_{\Omega} \varphi_i b \cdot \nabla \varphi_j d\Omega \\ &\quad + \left(\frac{1}{d\tau} + \frac{r}{2} \right) \int_{\Omega} \varphi_i \varphi_j d\Omega \end{aligned} \quad (5.16.2)$$

Instead of directly computing the integral over the whole domain, we can compute the integral element by element and then sum all the pieces together. We can scan all the triangles in the domain Ω and compute the integrals and matrices. There are only three basis functions are nonzero, which can be mapped into the reference triangle \hat{T} and evaluated over domain \hat{T} .

Taking use of the two equation:

$$\nabla \varphi_i = J^{-tr} \nabla \varphi_i \quad (5.17.1)$$

$$\frac{\partial F(\hat{v}, \hat{y})}{\partial (v, y)} = \det J \quad (5.17.2)$$

and introduce a method of approximation, we can decompose the RHS of equations in (5.16) into three important matrices:

$$\begin{aligned} P_1 &= \int_{\Omega} \nabla \varphi_i \cdot A \nabla \varphi_j d\Omega = \left[\int_T \nabla \varphi_i \cdot A \nabla \varphi_j dT \right]_{i,j=0,1,2} \\ &= |\det J| (J^{-tr} \hat{G})^{tr} \left(\int_{\hat{T}} A d\hat{T} \right) J^{-tr} \hat{G} \\ &\approx \frac{1}{2} |\det J| \hat{G}^{tr} J^{-1} \bar{A} J^{-tr} \hat{G} \end{aligned} \quad (5.18)$$

$$\begin{aligned} P_2 &= \int_{\Omega} \varphi_i b \cdot \nabla \varphi_j d\Omega \approx |\det J| \left[\int_{\hat{T}} \hat{\varphi}_i d\hat{T} \right]_{i=0,1,2} \bar{b}^{tr} J^{-tr} \hat{G} \\ &= |\det J| \frac{1}{6} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \bar{b}^{tr} J^{-tr} \hat{G} \end{aligned} \quad (5.19)$$

$$\begin{aligned} P_3 &= \left[\int_{\Omega} \varphi_i \varphi_j d\Omega \right]_{i,j=0,1,2} = |\det J| \left[\int_{\hat{T}} \hat{\varphi}_i \hat{\varphi}_j d\hat{T} \right]_{i,j=0,1,2} \\ &= |\det J| \frac{1}{24} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \end{aligned} \quad (5.20)$$

where

$$\hat{G} = |\nabla \hat{\phi}_i|_{i=0,1,2} = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \quad (5.21)$$

In (5.18-19), we set:

$$\bar{A} = \frac{1}{3} \sum_{i=0}^2 A_i, \quad \bar{b} = \frac{1}{3} \sum_{i=0}^2 b_i \quad (5.22)$$

In computation of local average \bar{A} and \bar{b} , which has the form of (4.8), we can use:

$$v \approx \bar{v} = \frac{1}{3} \sum_{i=0}^2 v_i \quad (5.23)$$

Note that in this approximation we take the average value of A and b of the reference triangle, which will make the algorithm more efficient and would provide sufficient accuracy when the resolution of computation domain is high enough.

Due to the use of local basis functions, each node can interact with at most, including itself, seven nodes. In order to save the memory of computation, sparse matrix need to be used.

5.7. Solution and Discussion

In order to check the correctness of this algorithm, the computational results are examined against the closed-form solution obtained from the special case. The default parameters are listed in the Table 5.1 and special case will be explicitly stated. We still take the vanilla European call and put option as example.

TABLE 5.1: Parameters

Initial Spot Price	$S_0 = 100$	Risk-free Rate	$r = 0.05$
Initial Volatility	$v_0 = 0.25$	Dividend Yield	$q = 0.01$
Mean Volatility	$\theta = 0.09$	Call Strike	$K_C = 110$
Mean Reversion Rate	$\kappa = 1.0$	Put Strike	$K_P = 90$
Volatility Variance	$\xi = 0.4$	Time to Maturity	$T = 1.0$
Correlation	$\rho = -0.7$		

And by original derivation of Rouah and Heston, we have the below closed form solution for European call option:

$$C(K) = e^x P_1 - K e^{-r\tau} P_2 \quad (5.24)$$

where

$$P_j = \mathbb{P}(\ln S_T > \ln K) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left[\frac{e^{-i\phi} f_j(\phi; x, v)}{i\phi} \right] d\phi, \quad j = 1, 2 \quad (5.25)$$

and $f_j(\phi; x, v)$ are the characteristic functions in Heston's model. Further details about this formula can be found in many literatures and textbooks.

In Figure 5.3 and 5.4, we present the solution surface of the European vanilla call option price and European vanilla put option price respectively for the model parameter listed in Table 5.1 with different temporal 500 and spatial resolution 500×500 . In acquiring all the results, we use bilinear interpolation because our domain here is rectangular and the mesh is structured.

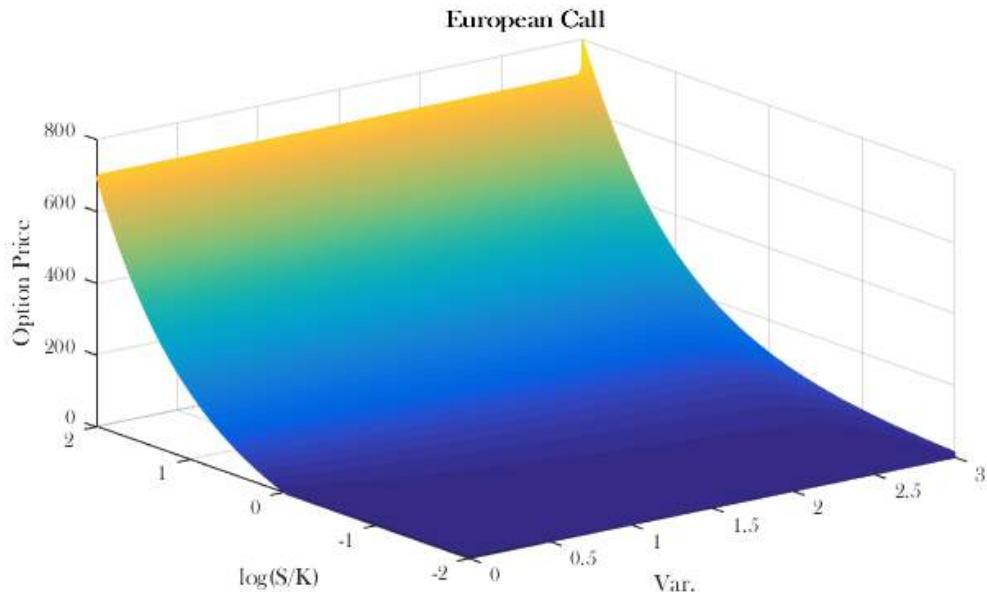


FIGURE 5.6: Solution Surface for European Call

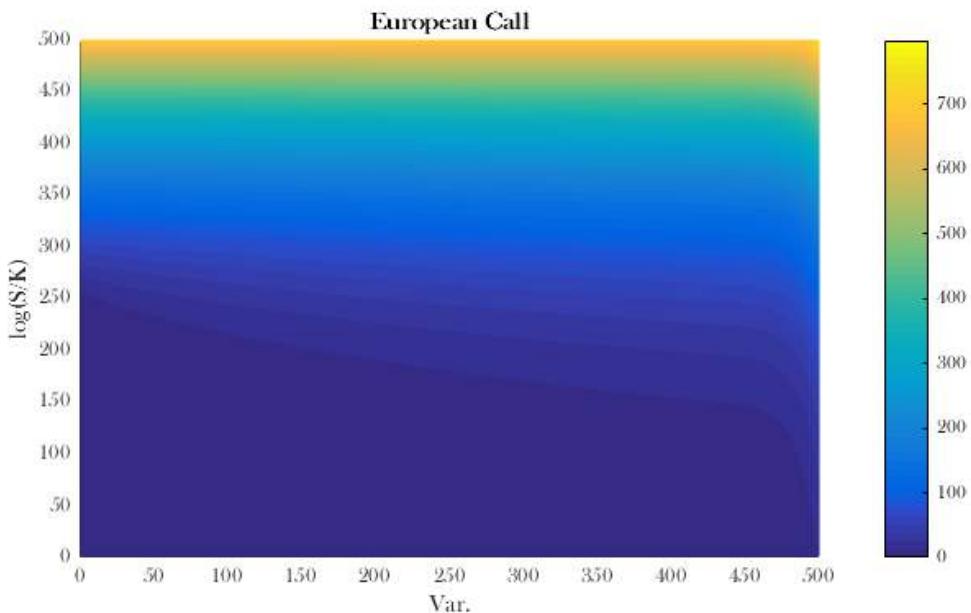
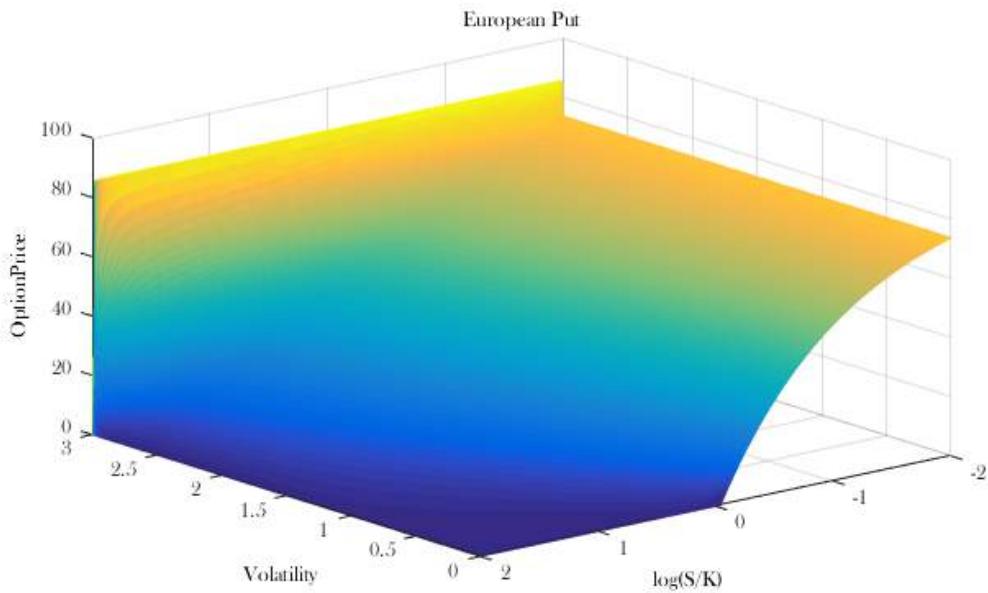
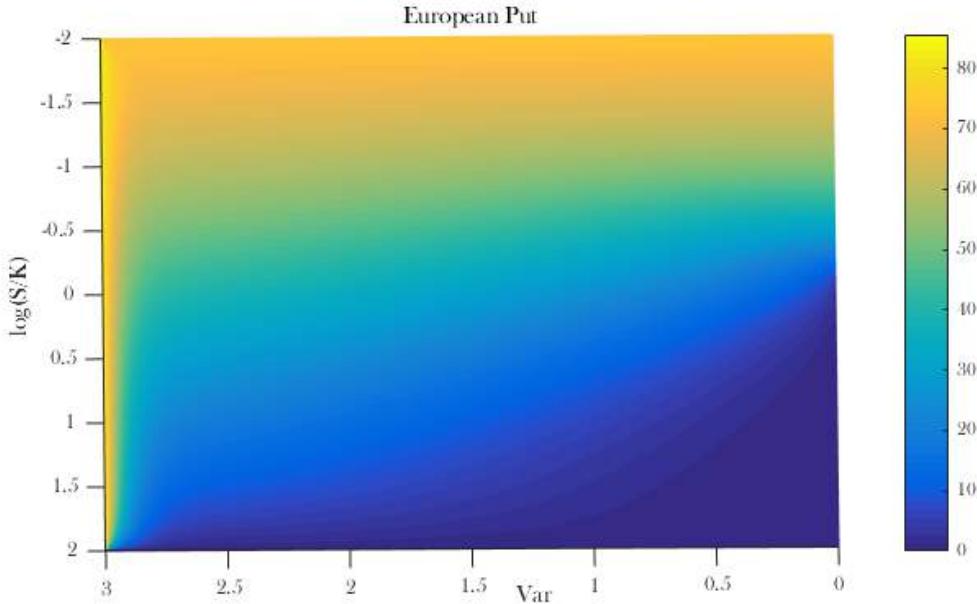


FIGURE 5.7: Solution Surface for European Call (Vertical View)

**FIGURE 5.8: Solution Surface for European Put****FIGURE 5.9: Solution Surface for European Put (Vertical View)**

In Table 5.2 and 5.3, we present the computational solution of FEM with Crank-Nicolson finite difference scheme for the European vanilla call option price and European vanilla put option price respectively, which are calibrated with the model parameter listed in Table 5.1 with different temporal and spatial resolution. And we also compare the computational solution with the closed-form solution to check the correctness of algorithm.

TABLE 5.2: European Vanilla Call Option Price Comparison

Resolution of (Variance, log(S/K), Time)	Computational Times(s)	Closed-Form Solution	FEM Computational Solution	Relative Error
(32, 32, 32)	4.355		13.7864	0.004829174
(64, 64, 64)	12.079		13.7717	0.005890293
(100, 100, 100)	37.834		13.7927	0.004374409
(200, 200, 200)	430.679	13.8533	13.8208	0.002346011
(300, 300, 300)	1830.4677		13.8318	0.001551977
(400, 400, 400)	5515.718		13.8376	0.001133304
(500, 500, 500)	13555.092		13.8412	0.000873438

TABLE 5.3: European Vanilla Put Option Price Comparison

Resolution of (Variance, log(S/K), Time)	Computational Times(s)	Closed-Form Solution	FEM Computational Solution	Relative Error
(32, 32, 32)	2.591		9.9944	0.007517304
(64, 64, 64)	11.641		10.0146	0.005511365
(100, 100, 100)	38.663		10.0331	0.003674244
(200, 200, 200)	407.649	10.7001	10.0512	0.001876843
(300, 300, 300)	1821.635		10.0575	0.001251229
(400, 400, 400)	5529.338		10.0606	0.00360706
(500, 500, 500)	14186.872		10.0625	0.000754709

Next we take consideration of European vanilla call option price for different strike price and maturity time with model parameter in Table 5.1. We choose spatial and temporal resolution as (100,100,100). The results are shown in Table 5.4 and 5.5.

TABLE 5.4: European Vanilla Call Option Price with different Strike Price

Strike Price	Closed-Form Solution	FEM Computational Solution	Relative Error
105	15.935692	15.8735	0.003902686
110	13.853335	13.7927	0.004376924
115	11.975267	11.9165	0.004907364
130	7.4758044	7.4225	0.007130256
150	3.6874922	3.6444	0.011686045

TABLE 5.5: European Vanilla Call Option Price with Different Maturity Time

Maturity Time T	Closed-Form Solution	FEM Computational Solution	Relative Error
1/12	2.1903889	2.1373	0.024237203
1/4	5.7981669	5.7427	0.009566282
1/2	9.3185161	9.2597	0.006311745
1	13.854340	13.7927	0.004449147

For European vanilla put option price for different strike price and maturity time with model parameter in Table 5.1, we also choose spatial and temporal resolution as (100,100,100). The results are shown in Table 5.6 and 5.7.

TABLE 5.6: European Vanilla Put Option Price with Different Strike Price

Strike Price	Closed-Form Solution	FEM Computational Solution	Relative Error
95	12.099818	12.0653	0.00285277
90	10.069421	10.0331	0.003607059
85	8.2562793	8.2191	0.004503154
80	6.6576145	6.6207	0.005544704
70	4.0796191	4.0462	0.008191721

TABLE 5.7: European Vanilla Put Option Price with Different Maturity Time

Maturity Time T	Closed-Form Solution	FEM Computational Solution	Relative Error
1/12	1.8359884	1.7913	0.024340241
1/4	4.7677247	4.7201	0.009988979
1/2	7.3419447	7.2974	0.006067153
1	10.069421	10.0331	0.003607059

6. Conclusion

This project has illustrated how to price Vanilla European option in a stochastic volatility model, such as Heston model, by using finite element method.

The results presented in this project shows that finite element method is an efficient and accurate method for solving option pricing problems, no matter univariate or multivariate problems. When taking use of Crank-Nicolson finite difference method scheme for temporal discretization, the numerical vanilla option values are fairly reliable and accurate, comparing with closed-form solutions.

Finite element method deserves our more attention for financial analysis and with appropriate definition of boundary condition, finite element method can be easily extended to price more complicated exotic options. It can also be further improved by applying adaptive mesh refinement and/or higher order basis functions.

However, the problem of checking boundary when assign value in the sparse matrices can result the longer computing time and using average values as approximations might bring error. As the relative errors of our solution are generally 10^{-3} and computation times are generally 10^4 seconds for high resolution, more adaptive work is needed for improvement in accuracy and efficiency.

References

- [1]. Achdou, Y., & Pironneau, O. (2007). Finite element methods for option pricing. Submitted for a special volume edited by R. Cont (Columbia UNY).
- [2]. Alberty, J., Carstensen, C., & Funken, S. A. (1999). Remarks around 50 lines of Matlab: short finite element implementation. Numerical Algorithms, 20(2-3), 117-137.
- [3]. Brenner, S., & Scott, R. (2007). The mathematical theory of finite element methods (Vol. 15). Springer Science & Business Media.
- [4]. Chen, L. O. N. G. (2011). Programming of finite element methods in MATLAB. Preprint, University of California Irvine, <http://math.uci.edu/~chenlong/226/Ch3FEMCode.pdf>.
- [5]. Chen, X., Burkardt, J., & Gunzburger, M. High Accuracy Finite Element Methods for Option Pricing under Heston's Stochastic Volatility Model.
- [6]. Duran, R. G. (1996). Galerkin approximations and finite element methods. Lecture notes (available at the author's website).
- [7]. Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of financial studies, 6(2), 327-343.
- [8]. LaForce, T. PE281 Boundary Element Method Course.
- [9]. Larcher, A., & Degirmenci, N. C. (2013). Lecture Notes: The Finite Element Method.
- [10]. Nikishkov, G. P. (2004). Introduction to the finite element method. University of Aizu.
- [11]. Reddy, J. N. (1989). An introduction to the finite element method.
- [12]. Süli, E. (2012). Lecture notes on Finite Element Methods for partial differential equations. Mathematical Institute, University of Oxford.
- [13]. Topper, J. (2000). Finite element modeling of exotic options. In Operations Research Proceedings 1999 (pp. 336-341). Springer Berlin Heidelberg.
- [14]. Topper, J. (2005). Option pricing with finite elements. Wilmott Magazine, 84-90.
- [15]. Våg, A. B. (2013). Pricing Put Options using Heston's Stochastic Volatility Model.
- [16]. Winkler, G., Apel, T., & Wystup, U. (2001). Valuation of options in Heston's stochastic volatility model using finite element methods. Foreign Exchange Risk, 283-303
- [17]. Xiong, C. (2015). Valuation of Double Barrier European Options in Heston's Stochastic Volatility Model Using Finite Element Methods (Doctoral dissertation, Rutgers University).
- [18]. Rouah, F. D. (2013). The Heston model and its extensions in MATLAB and C#. John Wiley & Sons.