Course Notes on Computational Finance

4

Finite-Difference Methods for American Vanilla Options

4. Finite-Difference Methods for American Vanilla Options

This chapter considers options with vanilla payoff $\Psi(S)$, where S is the price of an underlying asset. The assumed model for S_t is GBM. A continuous flow of dividend payments is admitted, with constant *dividend rate* $\delta \geq 0$. Hence the GBM is

$$\frac{\mathrm{d}S_t}{S_t} = (\mu - \delta)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t\,.$$

4.1 Preparations

The Black–Scholes equation for the value $V(S,t)$ of a standard option is

$$\frac{\partial V}{\partial t} + \frac{1}{2}\,\sigma^2\,S^2\,\frac{\partial^2 V}{\partial S^2} + (r-\delta)\,S\,\frac{\partial V}{\partial S} - r\,V = 0\,,$$

defined on the domain $S > 0$, $0 \leq t \leq T$. Terminal condition: $V(S,T) = \Psi(S)$.

This partial differential equation (PDE) can be solved directly with numerical methods. But then tricky stability issues must be tackled. Here we prefer applying transformations as much as possible in order to obtain simpler equations, e.g. with the transformation $S = \mathrm{e}^x$.

Assumption: σ, r and δ are constant.

Then the Black–Scholes equation can be transformed to a strikingly simple type of PDE (exercises). With $t = T - \frac{2\tau}{\sigma^2}$, $S = K\mathrm{e}^x$, $q := \frac{2r}{\sigma^2}$, $q_\delta := \frac{2(r-\delta)}{\sigma^2}$ and $y(x,\tau)$ defined by

$$V(S,t) = V(K\mathrm{e}^x, T - \frac{2\tau}{\sigma^2}) =: v(x,\tau)$$

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

one obtains

$$\boxed{\frac{\partial y}{\partial \tau} = \frac{\partial^2 y}{\partial x^2}}$$

This is a parabolic equation which has an analytic solution. A back transformation establishes the well-known Black-Scholes *formula*. So, for a standard option of European style no specific numerical method is needed.

The aim of the chapter is to price **American** options.

Generic applications of the simple PDE $y_\tau = y_{xx}$ are heat conduction and diffusion in a one-dimensional medium. The problem is *well-posed* as **initial-value problem** with increasing τ. The time transformation $t \to \tau$ converts the terminal condition for $V(S,T) = \Psi(S)$ to an initial condition for $\tau = 0$:

$$\text{call}: \quad y(x,0) = \max\left\{ \mathrm{e}^{\frac{x}{2}(q_\delta+1)} - \mathrm{e}^{\frac{x}{2}(q_\delta-1)}, 0 \right\}$$

$$\text{put}: \quad y(x,0) = \max\left\{ \mathrm{e}^{\frac{x}{2}(q_\delta-1)} - \mathrm{e}^{\frac{x}{2}(q_\delta+1)}, 0 \right\}$$

A solution y of the initial-value problem is defined on the domain

$$-\infty < x < +\infty, \quad 0 \le \tau \le \frac{1}{2}\sigma^2 T =: \tau_{\max}\,,$$

which is a strip in the (x,τ)-plane. "Boundary" here means $x \to -\infty$, $x \to +\infty$.

Boundary-Value Problem

For numerical purposes the infinite strip is truncated to a rectangle, with

$$x_{\min} \le x \le x_{\max} .$$

This "localization" cuts down the influence of the initial conditions, and additional conditions must be formulated to make the problem well-posed. These are boundary conditions for the sides $x_{\min}$ and $x_{\max}$. For the solution $w(x,\tau)$ of the **boundary-value problem** on the rectangle we aim at $w \approx y$, which requires the rectangle to be "large." The sides $x_{\min}$ and $x_{\max}$ must be chosen accordingly. In addition, the interval must include the range of financial interest, namely, the x-values of S_0 and K. This requires

$$x_{\min} < \min\left\{0,\, \log\frac{S_0}{K}\right\}, \quad \max\left\{0,\, \log\frac{S_0}{K}\right\} < x_{\max} .$$

For simplicity, just think of $x_{\min} = -3$ and $x_{\max} = 3$. — Later we shall see that for American options the partial differential equation mutates to an inequality.

4.2 Basics of Finite-Difference Methods (FDM)

A. Difference Approximations

Recall

$$f'(x) = \frac{f(x+h) - f(x)}{h} - \frac{h}{2} f''(\xi) \quad \text{for a } \xi \in (x, x+h) \text{ and } f \in C^2$$

We introduce an equidistant grid with grid points x_i

$$... < x_{i-1} < x_i < x_{i+1} < ...$$

and $h := x_{i+1} - x_i$. [In the non-equidistant case one often prefers finite-element methods (FEM)]. This chapter is confined to equidistant grids.

Analogously, with notation $f_i = f(x_i)$, the following holds true:

$$f'(x_i) = \frac{f_{i+1} - f_{i-1}}{2h} + O(h^2) \qquad \text{for } f \in C^3$$

$$f''(x_i) = \frac{f_{i+1} - 2f_i + f_{i-1}}{h^2} + O(h^2) \qquad \text{for } f \in C^4$$

$$f'(x_i) = \frac{-f_{i+2} + 4f_{i+1} - 3f_i}{2h} + O(h^2) \quad \text{for } f \in C^3$$

B. The Grid

For an $m \in \mathbb{N}$ and a $\nu_{\max} \in \mathbb{N}$ define

$$\Delta x := \frac{x_{\max} - x_{\min}}{m}\,, \quad x_i := x_{\min} + i \cdot \Delta x\,, \quad i = 0, 1, ..., m\,;$$
$$\Delta \tau := \frac{\tau_{\max}}{\nu_{\max}}\,, \quad \tau_\nu := \nu \cdot \Delta \tau\,, \quad \nu = 0, ..., \nu_{\max}\,.$$

$y_{i,\nu} := y(x_i, \tau_\nu)$ is the value of y at the node (x_i, τ_ν). Approximations of $y_{i,\nu}$ are denoted $w_{i,\nu}$. The values

$$\boxed{w_{i,0} = y(x_i, 0)}$$

are known from the initial conditions.

C. Explicit Method

In the PDE, replace

$$\frac{\partial y_{i,\nu}}{\partial \tau} := \frac{\partial y(x_i,\tau_\nu)}{\partial \tau} = \frac{y_{i,\nu+1}-y_{i,\nu}}{\Delta\tau} + O(\Delta\tau) \quad \text{and}$$
$$\frac{\partial^2 y_{i,\nu}}{\partial x^2} = \frac{y_{i+1,\nu}-2y_{i,\nu}+y_{i-1,\nu}}{\Delta x^2} + O(\Delta x^2)\,,$$

drop the O-error terms, replace $y \to w$, and obtain the *difference equation*

$$\frac{w_{i,\nu+1}-w_{i,\nu}}{\Delta\tau} = \frac{w_{i+1,\nu}-2w_{i,\nu}+w_{i-1,\nu}}{\Delta x^2}\,.$$

In case all values w are calculated for the time level ν, then the values of the time level $\nu+1$ are given by

$$w_{i,\nu+1} = w_{i,\nu} + \frac{\Delta\tau}{\Delta x^2}(w_{i+1,\nu}-2w_{i,\nu}+w_{i-1,\nu})\,.$$

With the notation

$$\lambda := \frac{\Delta\tau}{\Delta x^2}$$

this is rewritten as

$$\boxed{w_{i,\nu+1} = \lambda w_{i-1,\nu} + (1-2\lambda) w_{i,\nu} + \lambda w_{i+1,\nu}\,.} \qquad (*)$$

Start with $\nu = 0$, since there the $w_{i,0}$ are known. That is, for $\nu = 1$ all $w_{i,1}$ can be calculated with the explicit formula $(*)$, and similarly the $w_{i,\nu}$-values of the following time levels.

ν is the outer loop index and i the inner index. This suggests to describe the procedure by vectors and matrices: Use

$$w^{(\nu)} := (w_{1,\nu}, \ldots, w_{m-1,\nu})^{tr}$$

(so far, reasonable boundary components $w_{0,\nu}$ and $w_{m,\nu}$ are lacking) for the values of the ν-th time level, and the $(m-1) \times (m-1)$-matrix

$$A := A_{\text{expl}} := \begin{pmatrix} 1-2\lambda & \lambda & 0 & \cdots & 0 \\ \lambda & 1-2\lambda & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \lambda \\ 0 & \cdots & 0 & \lambda & 1-2\lambda \end{pmatrix}$$

Now the explicit method can be written

$$w^{(\nu+1)} = Aw^{(\nu)} \quad \text{for } \nu = 0, 1, 2, ..., \nu_{\max} - 1$$

up to a modification taking care of boundary conditions for $w_{0,\nu}$ and $w_{m,\nu}$. Preliminarily, for ease of presentation, we have set the boundary conditions to zero.

Example of Instability

$y_\tau = y_{xx}$ with $y(x,0) = \sin \pi x$, $x_0 = 0$, $x_m = 1$ and boundary conditions $= 0$.
Let us approximate $y(x = 0.2, \tau = 0.5)$ using a grid with $\Delta x = 0.1$, i.e. $m = 10$ and $0.2 = x_2$, and two different values of $\Delta\tau$:

a) $\Delta\tau = 0.0005 \quad \Rightarrow \quad \lambda = 0.05 \quad$ and $\quad 0.5 = \tau_{1000}$
yields $w_{2,1000} = 0.00435 \quad$ (exact is 0.004227)

b) $\Delta\tau = 0.01 \quad \Rightarrow \quad \lambda = 1 \quad$ and $\quad 0.5 = \tau_{50}$
yields $w_{2,50} = -1.5 \cdot 10^8 \quad$ (an instability)

D. Stability

Error analysis of $w^{(\nu+1)} = Aw^{(\nu)} + d^{(\nu)}$ (above $d^{(\nu)} = 0$)

Notation

$w^{(\nu)}$ exact vectors of $w^{(\nu+1)} = Aw^{(\nu)} + d^{(\nu)}$,

$\overline{w}^{(\nu)}$ versions calculated in the computer, subjected to rounding errors,

$e^{(\nu)} := \overline{w}^{(\nu)} - w^{(\nu)}$.

In exact computation, $\overline{w}^{(\nu+1)} - A\overline{w}^{(\nu)} - d^{(\nu)}$ is nonzero; we call this vector $r^{(\nu+1)}$,

$$\overline{w}^{(\nu+1)} = A\overline{w}^{(\nu)} + d^{(\nu)} + r^{(\nu+1)} .$$

The vector $r^{(\nu+1)}$ represents the rounding errors in the ν-th step.

It suffices to study the propagation of *one* error. So we set $r^{(\nu)} = 0$ for $\nu > 1$, i.e., study propagation of the error $e^{(0)}$ in the course of the iterations.

$\Rightarrow\ \overline{w}^{(\nu+1)} = A\overline{w}^{(\nu)} + d^{(\nu)} \quad (\nu > 1)$

$\Rightarrow\ Ae^{(\nu)} = A\overline{w}^{(\nu)} - Aw^{(\nu)} = \overline{w}^{(\nu+1)} - w^{(\nu+1)} = e^{(\nu+1)}$

$\Rightarrow\ e^{(\nu)} = A^{\nu}e^{(0)}$

For stable behavior, we require $A^{\nu}e^{(0)} \to 0$ for $\nu \to \infty$.

Notation:

$$\rho(A) := \max_k |\mu_k^A| \text{ where } \mu^A \text{ is eigenvalue of } A$$

Lemma 1

$$\rho(A) < 1 \iff A^\nu z \to 0 \text{ for all } z \text{ and } \nu \to \infty$$

Proof: Textbooks on Numerical Analysis.

A has the structure

$$A = I - \lambda \cdot \underbrace{\begin{pmatrix} 2 & -1 & & & 0 \\ -1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ 0 & & & -1 & 2 \end{pmatrix}}_{=:G}.$$

$\Rightarrow \mu^A = 1 - \lambda \mu^G$.

Lemma 2

$$\text{For } G = \begin{pmatrix} \alpha & \beta & & & 0 \\ \gamma & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \beta \\ 0 & & & \gamma & \alpha \end{pmatrix} \in \mathbb{R}^{N\times N} \text{ the eigenvalues are}$$

$$\mu_k^G = \alpha + 2\beta\sqrt{\frac{\gamma}{\beta}}\cos\frac{k\pi}{N+1}\ ,\ k = 1, ..., N.$$

Proof:

With the eigenvectors

$$v^{(k)} = \left(\sqrt{\frac{\gamma}{\beta}}\sin\frac{k\pi}{N+1}, \left(\sqrt{\frac{\gamma}{\beta}}\right)^2\sin\frac{2k\pi}{N+1}, ..., \left(\sqrt{\frac{\gamma}{\beta}}\right)^N\sin\frac{Nk\pi}{N+1}\right)^{tr}$$

check $Gv^{(k)} = \mu^G v^{(k)}$.

(For $\gamma = \beta$ the eigenvectors do not depend on α, β, γ.)

Applying Lemma 2 either directly to A, or to G with $N = m-1$, $\alpha = 2$ and $\beta = \gamma = -1$, yields

$$\mu_k^G = 2 - 2\cos\frac{k\pi}{m} = 4\sin^2\left(\frac{k\pi}{2m}\right)$$

$$\mu_k^A = 1 - 4\lambda\sin^2\frac{k\pi}{2m}\,.$$

By Lemma 1:

$$\text{stability} \iff \left|1 - 4\lambda\sin^2\frac{k\pi}{2m}\right| < 1, \quad k = 1, ..., m-1$$

$$\iff \lambda > 0 \text{ and } -1 < 1 - 4\lambda\sin^2\frac{k\pi}{2m}, \quad \text{or} \quad \frac{1}{2} > \lambda\sin^2\frac{k\pi}{2m}$$

For the second inequality $\lambda \leq 1/2$ is sufficient. In summary,

For $0 < \lambda \leq \frac{1}{2}$ the explicit method $w^{(\nu+1)} = Aw^{(\nu)}$ with $A = A_{\text{expl}}$ is stable.

Because of $\lambda = \frac{\Delta\tau}{\Delta x^2}$, or $0 < \Delta\tau \leq \frac{1}{2}\Delta x^2$, the step sizes $\Delta\tau$ and Δx can not be chosen independent of each other.

Conclusion: This explicit method is not satisfying.

E. An Implicit Method

The backward difference quotient

$$\frac{\partial y_{i,\nu}}{\partial \tau} = \frac{y_{i,\nu} - y_{i,\nu-1}}{\Delta\tau} + O(\Delta\tau).$$

leads to

$$\boxed{-\lambda w_{i+1,\nu} + (1+2\lambda) w_{i,\nu} - \lambda w_{i-1,\nu} = w_{i,\nu-1}\,.}$$

This is a system of coupled linear equations for the $w_{i,\nu}$. With

$$A := A_{\text{impl}} := \begin{pmatrix} 1+2\lambda & -\lambda & & 0 \\ -\lambda & \ddots & \ddots & \\ & \ddots & \ddots & -\lambda \\ 0 & & -\lambda & 1+2\lambda \end{pmatrix}$$

the vector $w^{(\nu)}$ is defined implicitly as solution of $Aw^{(\nu)} = w^{(\nu-1)}$, or

$$Aw^{(\nu+1)} = w^{(\nu)}\,,\ \nu = 0, \ldots, \nu_{\max} - 1.$$

(We still use the preliminary boundary conditions $w_{0,\nu} = w_{m,\nu} = 0$.) This method is called *backward difference method* or *backward time centered space (BTCS)* or *fully implicit*.

Stability

The above Lemmas imply that the method is *unconditionally stable*, and $\Delta\tau$ and Δx can be chosen independent of each other.

Costs

$\nu_{\max} O(m)$, since only one LR-decomposition of A for $\nu = 0$ is necessary (cheap for a tridiagonal matrix). For each ν, only one backward loop is required, which costs $O(m)$ operations.

A weakness of this method (and of the explicit method) is the accuracy of the first order in $\Delta\tau$, the error is of the order

$$O(\Delta x^2) + O(\Delta\tau)\,.$$

4.3 Crank–Nicolson Method

It would be desirable to have a stable method with error $O(\Delta\tau^2)$ for

$$\frac{\partial y}{\partial \tau} = \frac{\partial^2 y}{\partial x^2}\,.$$

From the previous section, we assemble the forward quotient for ν

$$\frac{w_{i,\nu+1} - w_{i,\nu}}{\Delta\tau} = \frac{w_{i+1,\nu} - 2w_{i,\nu} + w_{i-1,\nu}}{\Delta x^2}$$

and the backward quotient for $\nu + 1$

$$\frac{w_{i,\nu+1} - w_{i,\nu}}{\Delta\tau} = \frac{w_{i+1,\nu+1} - 2w_{i,\nu+1} + w_{i-1,\nu+1}}{\Delta x^2}\,.$$

Adding both equations yields the scheme

$$\frac{w_{i,\nu+1} - w_{i,\nu}}{\Delta\tau} = \frac{1}{2\Delta x^2}\left(w_{i+1,\nu} - 2w_{i,\nu} + w_{i-1,\nu} + w_{i+1,\nu+1} - 2w_{i,\nu+1} + w_{i-1,\nu+1}\right).$$

Theorem (Crank–Nicolson)

For this scheme the following assertions hold:

1.) For $y \in C^4$ the method is of the order $O(\Delta x^2) + O(\Delta\tau^2)$.

2.) For each ν a system of linear equations in tridiagonal form must be solved.

3.) The method is stable for all $\Delta\tau > 0$.

Proof:

1.) With the notation

$$\delta_{xx} w_{i,\nu} := \frac{w_{i+1,\nu} - 2w_{i,\nu} + w_{i-1,\nu}}{\Delta x^2}$$

the Taylor expansion for $y \in C^4$ yields

$$\delta_{xx} y_{i,\nu} = \frac{\partial^2}{\partial x^2} y_{i,\nu} + \frac{\Delta x^2}{12} \frac{\partial^4}{\partial x^4} y_{i,\nu} + O(\Delta x^4)\,.$$

Then the local discretization error

$$\varepsilon := \frac{y_{i,\nu+1} - y_{i,\nu}}{\Delta\tau} - \frac{1}{2}\left(\delta_{xx} y_{i,\nu} + \delta_{xx} y_{i,\nu+1}\right)$$

satisfies

$$\varepsilon = O(\Delta x^2) + O(\Delta\tau^2)\,.$$

2.) With $\lambda := \frac{\Delta\tau}{\Delta x^2}$ the scheme is rewritten

$$-\frac{\lambda}{2}w_{i-1,\nu+1} + (1+\lambda)w_{i,\nu+1} - \frac{\lambda}{2}w_{i+1,\nu+1}$$

$$= \frac{\lambda}{2}w_{i-1,\nu} + (1-\lambda)w_{i,\nu} + \frac{\lambda}{2}w_{i+1,\nu}\,.$$

With the preliminary boundary conditions $w_{0,\nu} = w_{m,\nu} = 0$ this is the matrix-vector system

$$Aw^{(\nu+1)} = Bw^{(\nu)}$$

where

$$A := \begin{pmatrix} 1+\lambda & -\frac{\lambda}{2} & & 0 \\ -\frac{\lambda}{2} & \ddots & \ddots & \\ & \ddots & \ddots & -\frac{\lambda}{2} \\ 0 & & -\frac{\lambda}{2} & 1+\lambda \end{pmatrix}, \quad B := \begin{pmatrix} 1-\lambda & \frac{\lambda}{2} & & 0 \\ \frac{\lambda}{2} & \ddots & \ddots & \\ & \ddots & \ddots & \frac{\lambda}{2} \\ 0 & & \frac{\lambda}{2} & 1-\lambda \end{pmatrix}.$$

By the theorem of Gerschgorin the eigenvalues of A lie between 1 and $1+2\lambda$. So, zero is no eigenvalue, A is non-singular, and the system of equations has a unique solution.

3.) Set

$$A = I + \tfrac{\lambda}{2}G \quad \text{with} \quad G := \begin{pmatrix} 2 & -1 & & 0 \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 2 \end{pmatrix} \quad \text{and} \quad B = I - \tfrac{\lambda}{2}G\,.$$

Then

$$\begin{aligned} \underbrace{(2I + \lambda G)}_{=:C} w^{(\nu+1)} &= (2I - \lambda G) w^{(\nu)} \\ &= (4I - 2I - \lambda G) w^{(\nu)} \\ &= (4I - C) w^{(\nu)}\,, \end{aligned}$$

which leads to the explicit form

$$w^{(\nu+1)} = (4C^{-1} - I) w^{(\nu)}.$$

By Lemma 1 the stability requirement is

$$\left| \frac{4}{\mu_k^C} - 1 \right| < 1 \quad \text{for all } k.$$

By Section 4.2D, the eigenvalues μ_k^C of C are

$$\mu_k^C = 2 + \lambda \mu_k^G = 2 + 4\lambda \sin^2 \frac{k\pi}{2m} \; > 2 \, .$$

Hence the method is stable for all $\lambda > 0$ ($\Delta\tau > 0$).

Algorithm (Crank–Nicolson)

start: Choose m, $\nu_{\max}$; calculate $\Delta x, \Delta\tau$

$w_i^{(0)} = y(x_i, 0) \quad (0 \le i \le m)$

LR-decomposition (or RL-decomposition) of A

loop: for $\nu = 0, 1, ..., \nu_{\max} - 1$:

$c := Bw^{(\nu)} + 0$ (preliminary boundary conditions 0)

Solve $Ax = c$ (using the LR/RL-decomposition)

$w^{(\nu+1)} := x$

4.4 Boundary Conditions

Since the initial conditions are active only for the interval $x_{\min} \le x \le x_{\max}$, boundary conditions

$y(x,\tau)$ for $x_{\min}$ and $x_{\max}$, or
$w_{0,\nu}$ and $w_{m,\nu}$ for $\nu = 1, ..., \nu_{\max}$

are needed to make the problem well-posed. These boundary conditions are artificial and have an influence on the accuracy of numerical solutions.

In the S-world, the GBM assumption for S_t implies

$$\begin{aligned} S(0) = 0 \quad &\Rightarrow \quad S_t = 0 \text{ for all } t \\ S(0) \to \infty \quad &\Rightarrow \quad S_t \text{ large for all } t \le T. \end{aligned}$$

Hence for all t

$$\begin{aligned} V_{\rm C}(S,t) &= 0 \quad \text{for } S = 0, \quad \text{and} \\ V_{\rm P}(S,t) &\to 0 \quad \text{for } S \to \infty\,. \end{aligned}$$

We use these boundary conditions as approximation also for $S_{\min} \approx 0$ and $S_{\max}$ large. And for the transformed x-values, provided $-x_{\min} = -x_0$ and $x_{\max} = x_m$ are large enough, as well:

$$\begin{array}{lll} \text{call: “left”} & (S=0) & V=0, \quad \text{i.e. } w_{0,\nu}=0 \text{ for all } \nu \\ \text{put: “right”} & (S\to\infty) & V=0, \quad \text{i.e. } w_{m,\nu}=0 \text{ for all } \nu \end{array}$$

So far, the argumentation holds for both European and American options. Now we turn to the question, what are reasonable V-values on the other ends of the interval?

We begin with **European** options, and apply the *put-call parity*

$$V_{\rm P} = V_{\rm C} - S\mathrm{e}^{-\delta(T-t)} + K\mathrm{e}^{-r(T-t)}\,.$$

This yields the boundary conditions at the other ends:

$$\begin{aligned} V_{\rm C}(S,t) &= S\mathrm{e}^{-\delta(T-t)} - K\mathrm{e}^{-r(T-t)} \quad \text{for } S\to\infty \\ V_{\rm P}(S,t) &= K\mathrm{e}^{-r(T-t)} - S\mathrm{e}^{-\delta(T-t)} \quad \text{for } S\to 0\,. \end{aligned}$$

After the transformation $(S,t)\to(x,\tau)$ the asymptotic behavior is:

$$\begin{array}{l} y(x,\tau)=r_1(x,\tau) \text{ for } x\to-\infty, \quad y(x,\tau)=r_2(x,\tau) \text{ for } x\to\infty, \text{ with} \\ \text{call:} \quad r_1(x,\tau):=0, \quad r_2(x,\tau):=\exp\left(\frac{1}{2}(q_\delta+1)x+\frac{1}{4}(q_\delta+1)^2\tau\right) \\ \text{put:} \quad r_1(x,\tau):=\exp\left(\frac{1}{2}(q_\delta-1)x+\frac{1}{4}(q_\delta-1)^2\tau\right), \quad r_2(x,\tau):=0 \end{array}$$

For the finite interval $a:=x_{\min}\le x\le x_{\max}=:b$, r_1, r_2 are the dominating terms of the boundary conditions and can be used as approximations. Accordingly, we choose

$$w_{0,\nu} = r_1(a, \tau_\nu)$$
$$w_{m,\nu} = r_2(b, \tau_\nu)\,.$$

These boundary conditions are of Dirichlet-type. In the Crank–Nicolson scheme the boundary conditions lead to the additional terms

$$-\frac{\lambda}{2}\, w_{0,\nu+1} = -\frac{\lambda}{2}\, r_1(a, \tau_{\nu+1})$$
$$\frac{\lambda}{2}\, w_{0,\nu} = \frac{\lambda}{2}\, r_1(a, \tau_\nu)$$

and

$$-\frac{\lambda}{2}\, w_{m,\nu+1} = -\frac{\lambda}{2}\, r_2(b, \tau_{\nu+1})$$
$$\frac{\lambda}{2}\, w_{m,\nu} = \frac{\lambda}{2}\, r_2(b, \tau_\nu),$$

which are represented by the vector

$$d^{(\nu)} := \frac{\lambda}{2} \cdot \begin{pmatrix} r_1(a, \tau_{\nu+1}) + r_1(a, \tau_\nu) \\ 0 \\ \vdots \\ 0 \\ r_2(b, \tau_{\nu+1}) + r_2(b, \tau_\nu) \end{pmatrix} = \frac{\lambda}{2} \cdot \begin{pmatrix} w_{0,\nu} + w_{0,\nu+1} \\ 0 \\ \vdots \\ 0 \\ w_{m,\nu} + w_{m,\nu+1} \end{pmatrix}$$

In the Crank–Nicolson algorithm the right-hand side of the system of equations is now

$$c := Bw^{(\nu)} + d^{(\nu)} ;$$

the matrix A is unchanged:

$$Aw^{(\nu)} = Bw^{(\nu)} + d^{(\nu)}$$

We still need to specify the boundary conditions (b.c.) for American-style options on the "other" side, namely, left-hand b.c. for the put and right-hand b.c. for the call.

4.5 Early-Exercise Structure — Free-Boundary Problems

For a better understanding of American options a further analysis of the solution structure is helpful.

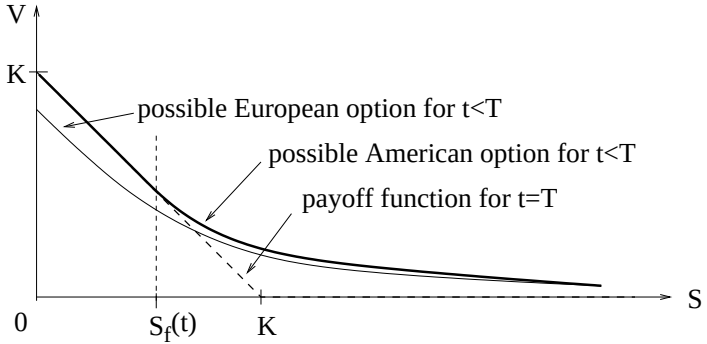
Basics (for the put under the Black–Scholes model)

a) $V(S,t)$ is continuous, $V \geq 0$, and $V \to 0$ for $S \to \infty$.

b) For $r > 0$ and all $t < T$

$$V_{\rm P}^{\rm Eu}(0,t) = K\mathrm{e}^{-r(T-t)} < K\,,$$

which implies: There is a $S^* = S^{*\rm Eu}(t)$ such that $V_{\rm P}^{\rm Eu}(S^*,t) = K - S^*$.



c) $V_{\rm P}$ is monotonic decreasing w.r.t. S, and convex. [R.C. Merton: Theory of Rational Option Pricing (1973)]

Hence S^* is unique.

d) $V_{\rm P}^{\rm Am} \geq (K-S)^+$ and $V_{\rm P}^{\rm Am} \leq K$

Assertion 1

Also for the American put with $r > 0$ there is an $S^* > 0$, such that $V_{\rm P}^{\rm Am}(S^*, t) = K - S^*$.

Proof:

Assume: $V_{\rm P}^{\rm Am} > K - S$ for all $S > 0$.
Then

$$-V_{\rm P}^{\rm Am} + K - S < 0\,,$$

i.e., exercising the put leads to a loss for all S. Accordingly, early exercise does not make sense, and hence $V_{\rm P}^{\rm Am} = V_{\rm P}^{\rm Eu}$. Consequently,

$$V_{\rm P}^{\rm Eu} = V_{\rm P}^{\rm Am} > K - S\,,$$

for all $S > 0$, which contradicts the implication of b).

Redefine $S^* := \max\{\, S \mid V^{\rm Am}(S,t) = K - S\,.\}$
By c) and d), $V = K - S$ for all $S < S^*$.

It remains to investigate the behavior of V for $S \geq S^*$.

Assertion 2

The right-hand side derivative of the function $V_{\mathrm{P}}^{\mathrm{Am}}(S,t)$ at S^* has the value -1.

Proof (outline):

We use the notation $V := V_{\mathrm{P}}^{\mathrm{Am}}$.

For the right-hand derivative $\frac{\partial V(S^*,t)}{\partial S} < -1$ is impossible, because otherwise for $S > S^*$ the property d) is violated. Hence $\frac{\partial V(S^*,t)}{\partial S} \geq -1$.

Assumption:

$$\frac{\partial V(S^*,t)}{\partial S} > -1\,.$$

We lead this to a contradiction as follows:

Build a portfolio: $\Pi := V + S$, with initial wealth

$$\Pi^* := V + S^*\,.$$

$(= K$; borrow from the bank the amount $K.)$ For our GBM

$$\mathrm{d}S = rS\,\mathrm{d}t + \sigma S\,\mathrm{d}W$$

assume that $\mathrm{d}t$ is so small that $\sqrt{\mathrm{d}t} \gg \mathrm{d}t$. If $\mathrm{d}t$ is small enough, then (intuitively)

$$\mathrm{d}W > 0 \quad \Longleftrightarrow \quad \mathrm{d}S > 0\,.$$

The Itô-Lemma leads to

$$\begin{aligned}\mathrm{d}\Pi &= (\ldots)\mathrm{d}t + \frac{\partial \Pi}{\partial S}\sigma S\,\mathrm{d}W \\ &= O(\mathrm{d}t) + \left(\frac{\partial V}{\partial S} + 1\right)\sigma S\,\mathrm{d}W.\end{aligned}$$

For $\mathrm{d}S > 0$ this is positive for sufficiently small $\mathrm{d}t$ because $\mathrm{d}W > 0$.
For $\mathrm{d}S < 0$ the wealth of the portfolio is $\Pi \equiv K$ and hence $\mathrm{d}\Pi = 0$.

In summary: $\mathsf{E}(\mathrm{d}\Pi) > 0$, and $\mathsf{E}(\mathrm{d}\Pi)$ is of the order $O(\sqrt{\mathrm{d}t})$. Sell the portfolio after $\mathrm{d}t$ and expect the following balance

$$\begin{aligned}\Pi^* + \mathsf{E}(\mathrm{d}\Pi) - K\mathrm{e}^{r\mathrm{d}t} &= V + S^* + \mathsf{E}(\mathrm{d}\Pi) - K(1 + O(\mathrm{d}t)) \\ &= \mathsf{E}(\mathrm{d}\Pi) + O(\mathrm{d}t)\end{aligned}$$

This is positive because $\mathsf{E}(\mathrm{d}\Pi)$ dominates, and we have arrived at a contradiction to the no-arbitrage principle.

Often related proofs use an argument of maximizing the value of the option. In this way, the *perpetual option* (an option that does not expire) can be analyzed, see the Exercises.

In summary, for an American put the following holds:

(1) $V(S^*, t) = K - S^*$ (Dirichlet b.c.)
(2) $\frac{\partial V(S^*,t)}{\partial S} = -1$ (Neumann b.c.)
(3) $V(S, t) > K - S$ for $S > S^*$
(4) $V(S, t) = K - S$ for $S \leq S^*$

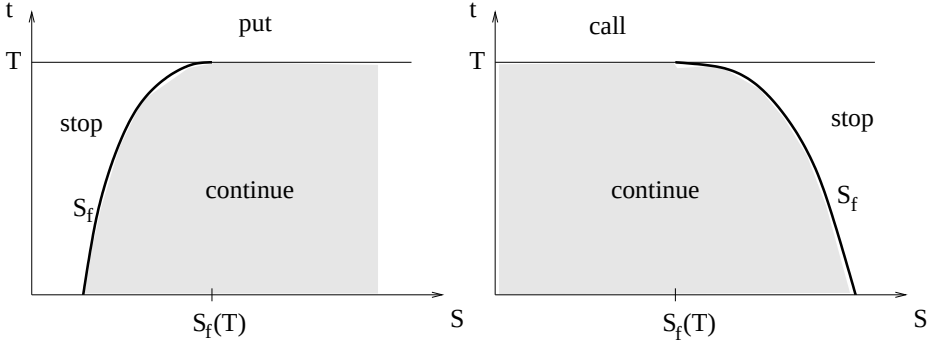
The property (2) is called *high contact*, or *smooth pasting*.

Such an S^* exists for each t. This defines a function, which we denote $S_{\rm f}(t)$. The "f" stands for *free boundary*.

The curve $S_{\rm f}(t)$ cuts the half strip into two parts, namely,

1.) $S > S_{\rm f}$, called *continuation region* of the put.

2.) $S \leq S_{\rm f}$, called *stopping region* of the put.



For standard options without dividend payments, these domains are simply connected. The curve $S_{\mathrm{f}}(t)$ is the interface.

Early-Exercise Curve

The curve $S_{\mathrm{f}}(t)$ is the *early-exercise curve* by the following reasons:

1.) In case $V_{\mathrm{P}} > (K-S)^+$, exercising amounts to $-V+K-S<0$. This is a loss. Consequently, the holder continues to *hold* the option.

2.) In case the price S passes the curve, $S < S_{\mathrm{f}}(t)$, then immediate exercising makes sense ("stopping"), because the amount K can be invested, leading for $r>0$, $t<T$ to the profit:

$$K\mathrm{e}^{r(T-t)} - K = K(\mathrm{e}^{r(T-t)} - 1).$$

By exercising, the final balance $K\mathrm{e}^{r(T-t)}$ is larger than $S\mathrm{e}^{\delta(T-t)}$, at least for $r(T-t)<1$.

Free-Boundary Problem means:

The Black–Scholes equation is valid only in the continuation region, not in the stopping region. Hence the domain for the BS equation for an American-style put is

$$S_{\mathrm{f}}(t) < S < \infty\,.$$

The left-hand boundary $S_{\mathrm{f}}(t)$ is "free" in the sense that it is unknown initially. It is calculated numerically, based on the additional boundary condition provided by the contact condition $\frac{\partial V}{\partial S} = -1$. This condition fixes the location of $S_{\mathrm{f}}(t)$.

The properties of a call are derived analogously.

Important Properties of the early-exercise curve in case **of a put** under the Black–Scholes model (continuous dividend rate $\delta \geq 0$ allowed, but discrete dividend here excluded) are:

1.) $S_{\mathrm{f}}(t)$ is continuously differentiable for $t < T$.

2.) $S_{\mathrm{f}}(t)$ is monotonic increasing.

3.)

$$S_{\mathrm{f}}(T) := \lim_{\substack{t\to T\\ t<T}} S_{\mathrm{f}}(t) = \begin{cases} K & \text{für } 0 \leq \delta \leq r \\ \frac{r}{\delta} K & \text{für } r < \delta \end{cases}$$

Proof of 3.) (notation $V = V_{\mathrm{P}}^{\mathrm{Am}}$)

$V^{\mathrm{Am}} \geq V^{\mathrm{Eu}}$ implies $S_{\mathrm{f}}(t) \leq S^{*\mathrm{eu}}(t) < K$ for all $t < T$. Hence $S_{\mathrm{f}}(T) \leq K$.

Notice that $V \geq \Psi$ implies

$$\frac{\partial V(S,t)}{\partial t} \leq 0$$

for $t = T$. To prepare for some indirect proofs, we first study for $S_{\mathrm{f}}(T) < K$ how the BS equation is consistent with the sign of $\frac{\partial V}{\partial t}$. At $t = T$ and for $S_{\mathrm{f}}(T) < S < K$

$$V(S,T) = K - S$$

holds, and the BS equation is

$$\frac{\partial V}{\partial t} + 0 - (r-\delta)S - rV = 0$$

$$\Longrightarrow \quad \frac{\partial V(S,T)}{\partial t} = rK - \delta S\,.$$

We will check for which combinations of (δ, r) the sign of $\frac{\partial V(S,T)}{\partial t}$ is consistent with $\frac{\partial V}{\partial t} \leq 0$.

Case $\delta > r$

Here $\frac{r}{\delta}K < K$. Then either $S_{\mathrm{f}}(T) = \frac{r}{\delta}K$ (the assertion), or there exists one of two open intervals (i) $S_{\mathrm{f}}(T) < \frac{r}{\delta}K$ and (ii) $\frac{r}{\delta}K < S_{\mathrm{f}}(T)$.

(i) For S in the interval $S_{\mathrm{f}}(T) < S < \frac{r}{\delta}K$ we have $\frac{\partial V(S,T)}{\partial t} = rK - \delta S > 0$, a contradiction to

$$\frac{\partial V}{\partial t} \leq 0\,.$$

(ii) For S in the interval $\frac{r}{\delta}K < S < S_{\mathrm{f}}(T)$ there is a small $\mathrm{d}t$ such that $(S, T - \mathrm{d}t)$ is in the stopping area. The inequality $rK < \delta S$ holds, and thus

$$K(\mathrm{e}^{r\mathrm{d}t} - 1) < S(\mathrm{e}^{\delta \mathrm{d}t} - 1)\,,$$

since $S < K$. This means that in case of stopping the dividend yield is larger than the return of investing K at the risk-free rate. Hence early exercise is not optimal, which is in conflict with the meaning of $S < S_{\mathrm{f}}(t)$.

Hence $S_{\mathrm{f}}(T) = \frac{r}{\delta}K$ holds in case $\delta > r$.

Case $\delta \leq r$

Assume $S_{\mathrm{f}}(T) < K$. Then for S in the interval $S_{\mathrm{f}}(T) < S < K$ a contradiction is obtained from

$$\underbrace{\frac{\partial V}{\partial t}}_{\leq 0} = \underbrace{rK - \delta S}_{>0} .$$

For a call with $\delta > 0$ but without discrete dividend:

1.) $S_{\mathrm{f}}(t)$ is continuously differentiable for $t < T$.

2.) $S_{\mathrm{f}}(t)$ is monotonic decreasing.

3.)

$$S_{\mathrm{f}}(T) := \lim_{\substack{t \to T \\ t<T}} S_{\mathrm{f}}(t) = \max\left\{K, \frac{r}{\delta}K\right\}$$

Remark: In case of discrete dividend payment the above assertions must be modified. In particular, S_{f} then is not continuous! For example, for an American put early exercise is not optimal within a certain time interval before ex-dividend date.

4.6 Linear Complementarity

We need a numerical method that does not use the unknown S_f explicitly.

A. Inequality

As argued above, for American options the Black–Scholes equation does not hold in the entire half strip $0 \le t \le T,\ 0 < S$, but only in the continuation region. Now, what happens in the stopping region?

With the notation:

$$\mathcal{L}_{\mathrm{BS}}(V) := \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r-\delta) S \frac{\partial V}{\partial S} - rV$$

the Black–Scholes equation can be written

$$\frac{\partial V}{\partial t} + \mathcal{L}_{\mathrm{BS}}(V) = 0\,.$$

For for a put and $S < S_\mathrm{f}$ the identity $V \equiv \Psi$ holds, so

$$V = K - S, \quad \frac{\partial V}{\partial S} = -1\,, \quad \frac{\partial^2 V}{\partial S^2} = 0\,, \quad \frac{\partial V}{\partial t} = 0\,,$$

hence

$$\frac{\partial V}{\partial t} + \mathcal{L}_{\rm BS}(V) = \delta S - rK\,.$$

For $S < S_{\rm f}(T)$ we conclude

in case $r < \delta$: $S < \frac{r}{\delta}K$, hence $\delta S - rK < 0$

in case $r \geq \delta$: $\delta S - rK \leq r(S-K) < 0$ because $S < K$

and thus

$$\frac{\partial V}{\partial t} + \mathcal{L}_{\rm BS}(V) < 0\,.$$

For a call the same result follows by analogous arguments.

In summary for all $S > 0,\ 0 \leq t \leq T$ the partial **differential inequality**

$$\frac{\partial V}{\partial t} + \mathcal{L}_{\rm BS}(V) \leq 0$$

holds, both for put and call.

Overview

put: $V_{\mathrm{P}}^{\mathrm{Am}} = K - S \quad$ for $S \le S_{\mathrm{f}}$

$V_{\mathrm{P}}^{\mathrm{Am}}$ solves the BS equation for $S > S_{\mathrm{f}}$

contact condition: $\dfrac{\partial V(S_{\mathrm{f}}, t)}{\partial S} = -1$

call: $V_{\mathrm{C}}^{\mathrm{Am}} = S - K \quad$ for $S \ge S_{\mathrm{f}}$

$V_{\mathrm{C}}^{\mathrm{Am}}$ solves the BS equation for $S < S_{\mathrm{f}}$

contact condition: $\dfrac{\partial V(S_{\mathrm{f}}, t)}{\partial S} = 1$

The second derivative of V with respect to S is not continuous at S_{f}. That is, the value function V is smooth in the interior of the continuation region, but not on the entire half strip.

Remark: The transformation of Section 4.1 leads to

$$\frac{\partial V}{\partial t} + \mathcal{L}_{\mathrm{BS}}(V) = -\frac{\partial y}{\partial \tau} + \frac{\partial^2 y}{\partial x^2}\,.$$

B. Formulation with Penalty Term

A unified treatment of $\frac{\partial V}{\partial t} + \mathcal{L}_{\mathrm{BS}}(V) \le 0$ on the entire half strip is possible. To this end, introduce a suitable function $p(V) \ge 0$ requiring the penalty PDE

$$\frac{\partial V}{\partial t} + \mathcal{L}_{\mathrm{BS}}(V) + p(V) = 0$$

to hold. The *penalty term* p should be 0 in the continuation region, and positive in the stopping region. The distance to S_{f} is not known, but the distance $V - \Psi$ of V to the payoff Ψ is available and is used as control. One example of a penalty function is

$$p(V) := \frac{\epsilon}{V - \Psi} \quad \text{for a small } \epsilon > 0\,.$$

Let $V_\epsilon(S,t)$ denote the solution of the penalty PDE. Two extreme cases characterize the effect of the penalty term for (S,t) in the continuation area and in the stopping area:

- $V_\epsilon - \Psi \gg \epsilon$ implies $p \approx 0$. Then essentially the Black–Scholes equation results.
- $0 < V_\epsilon - \Psi \ll \epsilon$ implies a large value of p, which means that the BS-part of the equation is dominated by p. The BS equation is switched off, and $V_\epsilon \approx \Psi$.

The corresponding branches of the solution V_ϵ may be called the "continuation branch" ($p \approx 0$) and the "stopping branch" ($V_\epsilon \approx \Psi$). Obviously these two branches approximate the true solution V of the Black–Scholes problem. The intermediate range $V_\epsilon - \Psi \approx O(\epsilon)$ characterizes a boundary layer between the continuation branch and the stopping branch. In this layer around the early-exercise curve S_f the solution V_ϵ can be seen as a connection between the BS surface and the payoff plane.[a]

Remarks: p and the resulting PDE are nonlinear in V. An implementation that avoids $V_\epsilon \leq \Psi$ is not easy; not every choice of ϵ or Δt will be successful.

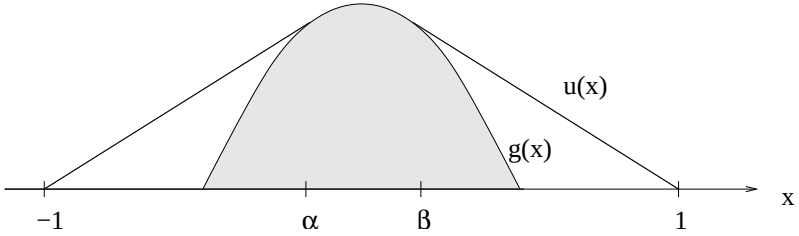
Penalty methods are powerful in general. But for the relatively simple situation of the single-asset American option, a more elegant solution is possible. We shall describe this approach next.

[a] This is illustrated in Topic 9, see the *Topics for CF* on the homepage `www.compfin.de`; Topic 9 also also illustrates the penalty function p.

C. Simple Obstacle Problem

Consider an "obstacle" g satisfying[b]

$$g > 0 \text{ on a subinterval of } -1 < x < 1,\ g \in C^2,\ g'' \le 0,\ g(-1) < 0 \text{ and } g(1) < 0.$$



A function $u \in C^1$ with minimal length, and with $u(-1) = u(1) = 0$ and $u \ge g$ for the g of the figure can be characterized as follows: There is α, β such that

$$\begin{aligned} -1 < x < \alpha: &\quad u'' = 0 \quad (u > g) \\ \alpha < x < \beta: &\quad u = g \quad (u'' = g'' \le 0) \\ \beta < x < 1: &\quad u'' = 0 \quad (u > g) \end{aligned}$$

This expresses a complementarity in the sense of

[b] This and other parts of this chapter were inspired by [P. Wilmott, J. Dewynne, S. Howison: The Mathematics of Financial Derivatives].

$$\text{in case } u-g>0, \text{ then } u''=0,$$
$$\text{in case } u-g=0, \text{ then } u''\le 0.$$

For options, the analogy is

$$\text{in case } V-\Psi>0, \text{ then } \dot{V}+\mathcal{L}_{\mathrm{BS}}(V)=0\,,$$
$$\text{in case } V-\Psi=0, \text{ then } \dot{V}+\mathcal{L}_{\mathrm{BS}}(V)\le 0\,.$$

For the simple obstacle problem an equivalent formulation is

Find a function u, such that
$$u''(u-g)=0, \quad -u''\ge 0, \quad u-g\ge 0, \qquad (*)$$
$$u(-1)=u(1)=0,\ u\in\mathcal{C}^1[-1,1].$$

In this version, α and β do not occur explicitly. After the numerical solution, the values of α and β will become apparent. The problem $(*)$ is a **linear complementarity problem** (LCP); it has the form

$$\mathcal{A}\cdot\mathcal{B}=0,\ \mathcal{A}\ge 0,\ \mathcal{B}\ge 0, \qquad \text{for suitable } \mathcal{A},\mathcal{B}.$$

For a **numerical solution** of the simple obstacle problem we use the grid

$$x_i = -1 + i\Delta x, \quad \Delta x := \frac{2}{m}, \quad g_i := g(x_i)\,.$$

This leads to the discrete form of the obstacle problem

$$\left.\begin{aligned}(w_{i-1} - 2w_i + w_{i+1})(w_i - g_i) &= 0,\\ -w_{i-1} + 2w_i - w_{i+1} &\geq 0,\\ w_i - g_i &\geq 0\end{aligned}\right\} \quad \text{for } 0 < i < m$$

with $w_0 = w_m = 0$. Correspondingly we define

$$G := \begin{pmatrix} 2 & -1 & & 0 \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 2 \end{pmatrix} \quad \text{and} \quad w := \begin{pmatrix} w_1 \\ \vdots \\ w_{m-1} \end{pmatrix}, \; g := \begin{pmatrix} g_1 \\ \vdots \\ g_{m-1} \end{pmatrix}$$

to rewrite the linear complementarity problem in vector notation:

$$\left.\begin{aligned}(w-g)^{tr} G w &= 0\\ Gw &\geq 0\\ w - g &\geq 0\end{aligned}\right\}$$

D. Complementarity of the Black–Scholes Problem

Specifically for the American put let us formulate the transformed problem

$$\frac{\partial y}{\partial \tau} = \frac{\partial^2 y}{\partial x^2} \quad \text{as long as} \quad V_{\mathrm{P}}^{\mathrm{Am}} > (K-S)^+$$

as LCP. This requires to transform also the side conditions:

$$\begin{aligned} V_{\mathrm{P}}^{\mathrm{Am}}(S,t) &\geq (K-S)^+ \\ &= K \max\{1-\mathrm{e}^x, 0\} \end{aligned}$$

with $q = \frac{2r}{\sigma^2}$ leads to (here specially for $\delta = 0$):

$$\begin{aligned} y(x,\tau) &\geq \exp\{\tfrac{1}{2}(q-1)x + \tfrac{1}{4}(q+1)^2\tau\} \max\{1-\mathrm{e}^x,\, 0\} \\ &= \exp\{\tfrac{1}{4}(q+1)^2\tau\} \max\{(1-\mathrm{e}^x)\mathrm{e}^{\frac{1}{2}(q-1)x},\, 0\} \\ &= \exp\{\tfrac{1}{4}(q+1)^2\tau\} \max\{\mathrm{e}^{\frac{1}{2}(q-1)x} - \mathrm{e}^{\frac{1}{2}(q+1)x},\, 0\} \\ &=: g(x,\tau). \end{aligned}$$

It turns out that both the boundary- and the initial conditions can be written with this g:

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

Now the linear complementarity problem for the American put is written

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

with boundary- and initial conditions as stated above.

This formulation is identical to that one for the general case $\delta \neq 0$, except for an adapted g:

$$q = \frac{2r}{\sigma^2}\,, \quad q_\delta = \frac{2(r-\delta)}{\sigma^2}\,,$$

put: $(r > 0)$ $\quad g(x,\tau) := \exp\{\frac{1}{4}((q_\delta - 1)^2 + 4q)\tau\} \max\{\mathrm{e}^{\frac{1}{2}(q_\delta - 1)x} - \mathrm{e}^{\frac{1}{2}(q_\delta + 1)x},\, 0\}$

call: $(\delta > 0)$ $\quad g(x,\tau) := \exp\{\frac{1}{4}((q_\delta - 1)^2 + 4q)\tau\} \max\{\mathrm{e}^{\frac{1}{2}(q_\delta + 1)x} - \mathrm{e}^{\frac{1}{2}(q_\delta - 1)x},\, 0\}$

Remark: For an American call without dividend and $r > 0$, $t < T$, the equality $V_{\mathrm{C}}^{\mathrm{Am}} = V_{\mathrm{C}}^{\mathrm{Eur}}$ holds. *Proof* for $\delta = 0$:

$$V_{\mathrm{C}}^{\mathrm{Am}} \geq V_{\mathrm{C}}^{\mathrm{Eur}} \geq S - K\mathrm{e}^{-r(T-t)} > S - K$$

Hence $V_{\mathrm{C}}^{\mathrm{Am}}$ >payoff; no early exercise.

4.7 Numerical Realization

A. Discretization

Now we discretize the LCP-problem with the grid from Section 4.2B:

$$\frac{w_{i,\nu+1}-w_{i,\nu}}{\Delta\tau} = \theta\frac{w_{i+1,\nu+1}-2w_{i,\nu+1}+w_{i-1,\nu+1}}{\Delta x^2} + (1-\theta)\frac{w_{i+1,\nu}-2w_{i,\nu}+w_{i-1,\nu}}{\Delta x^2}$$

("theta method"), with $\theta = 0$ (explicit method), $\theta = \frac{1}{2}$ (Crank–Nicolson) or $\theta = 1$ (BTCS-method). With the notation $\lambda := \frac{\Delta\tau}{(\Delta x)^2}$ the inequality $\frac{\partial y}{\partial\tau} - \frac{\partial^2 y}{\partial x^2} \geq 0$ becomes

$$\begin{aligned} w_{i,\nu+1} &- \lambda\theta(w_{i+1,\nu+1}-2w_{i,\nu+1}+w_{i-1,\nu+1}) \\ &- w_{i,\nu} - \lambda(1-\theta)(w_{i+1,\nu}-2w_{i,\nu}+w_{i-1,\nu}) \geq 0\,. \end{aligned} \qquad (*)$$

Ordering these terms leads to define for the ν-level terms

$$b_{i,\nu} := w_{i,\nu} + \lambda(1-\theta)(w_{i+1,\nu}-2w_{i,\nu}+w_{i-1,\nu}),\ \text{for } i = 2,\ldots,m-2.$$

The boundary conditions are included in $b_{1,\nu}$ and $b_{m-1,\nu}$:

$$b_{1,\nu} := w_{1,\nu} + \lambda(1-\theta)(w_{2,\nu} - 2w_{1,\nu} + g_{0,\nu}) + \lambda\theta g_{0,\nu+1}$$
$$b_{m-1,\nu} := w_{m-1,\nu} + \lambda(1-\theta)(g_{m,\nu} - 2w_{m-1,\nu} + w_{m-2,\nu}) + \lambda\theta g_{m,\nu+1}\,,$$

where

$$g_{i,\nu} := g(x_i, \tau_\nu) \quad (0 \le i \le m,\ 0 \le \nu \le \nu_{\max}).$$

This completes the vector

$$b^{(\nu)} := (b_{1,\nu}, \ldots, b_{m-1,\nu})^{tr}$$

and analogously, $w^{(\nu)}, g^{(\nu)}$. With the matrix

$$A := \begin{pmatrix} 1+2\lambda\theta & -\lambda\theta & & 0 \\ -\lambda\theta & \ddots & \ddots & \\ & \ddots & \ddots & -\lambda\theta \\ 0 & & -\lambda\theta & 1+2\lambda\theta \end{pmatrix} \in \mathbb{R}^{(m-1)\times(m-1)}$$

the problem $(*)$ is reformulated as

$$Aw^{(\nu+1)} \ge b^{(\nu)} \quad \text{for all } \nu.$$

And $y - g \geq 0$ reads

$$w^{(\nu)} \geq g^{(\nu)},$$

and finally $\left(\frac{\partial y}{\partial \tau} - \frac{\partial^2 y}{\partial x^2}\right)(y - g) = 0$ becomes

$$\left(Aw^{(\nu+1)} - b^{(\nu)}\right)^{tr} \left(w^{(\nu+1)} - g^{(\nu+1)}\right) = 0.$$

This constitutes the following macro-algorithm:

Algorithm

For $\nu = 0, 1, ..., \nu_{\max} - 1$:
compute $g := g^{(\nu+1)}$, $b := b^{(\nu)}$, as above;
compute w as solution of
$Aw - b \geq 0, \quad w \geq g, \quad (Aw - b)^{tr}(w - g) = 0.$ $(**)$
set $w^{(\nu+1)} := w$

For each time level ν the LCP $(**)$ must be solved. This topic is analyzed next.

Assertion

With the transformation $x := w - g$ and $y := Aw - b$ the LCP $(**)$ is equivalent to

> Compute vectors x, y, such that for $\hat{b} := b - Ag$ the following holds:
> $Ax - y = \hat{b}, \quad x \geq 0, \quad y \geq 0, \quad x^{tr} y = 0$

$(***)$

Notice that in this context x, y are general vectors $\in \mathbb{R}^{m-1}$.

Proof: Apply the transformation.

Lemma

The problem $(***)$ has a unique solution.

Proof:

1.) Define

$$G(x) := \frac{1}{2} x^{tr} A x - \hat{b}^{tr} x$$

with the A from above,

$$A = \mathrm{I} + \lambda\theta \begin{pmatrix} 2 & -1 & & 0 \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 2 \end{pmatrix}$$

By Lemma 2 of Section 4.2D the eigenvalues of the matrix A are

$$\mu_k^A = 1 + \lambda\theta 4 \sin^2(k\pi/2m) \geq 1, \quad k = 1, \ldots, m-1\,.$$

So the symmetric matrix A is positive definite. Differentiating G w.r.t. x yields

$$G_x = Ax - \hat{b}\,, \quad G_{xx} = A\,,$$

which shows that A is the Hessian matrix of G. For $G \in C^2$ recall

G is strictly convex. $\Longleftrightarrow$ The Hessian matrix of G is positive definite.

Hence the quadratic form G defined above is strictly convex and has a unique minimum on each convex subset of $\mathbb{R}^N$, $N = m-1$. Notice that $x \geq 0$ defines a convex set.

2.) Apply the Theorem of Karush, Kuhn, Tucker (KKT):

For the minimum x^* of a convex function G under the side conditions $H_i(x) \le d_i$ for $i = 1, \ldots, N$, where H_i are convex, the following holds: There are $y_i \ge 0$ such that

$$\frac{\partial G(x^*)}{\partial x_j} + y_1 \frac{\partial H_1(x^*)}{\partial x_j} + \ldots + y_N \frac{\partial H_N(x^*)}{\partial x_j} = 0 \quad (j = 1, \ldots, N)$$

with $H_i(x^*) \le d_i$ and $y_i(H_i(x^*) - d_i) = 0$ for $i = 1, \ldots, N$.

(references include [Stoer & Witzgall], [Strang])

In our application, $N = m - 1$. Inequality $x \ge 0$ or $x_i \ge 0$ for all $i = 1, \ldots, m-1$ leads to $d_i = 0$ and $H_i(x) = -x_i$. The Theorem of KKT implies the existence of $y \ge 0$ with

$$\frac{\partial G}{\partial x_j} + y_j \frac{\partial H_j}{\partial x_j} + 0 = \frac{\partial G}{\partial x_j} - y_j = 0$$

and $y_i x_i = 0$ for all i. For our special G we conclude

$$G_x = Ax - \hat{b} \quad \Longrightarrow \quad Ax - \hat{b} - y = 0$$

Hence $(***)$ results from the KKT-Theorem and processes a unique solution. This then carries over to $(**)$.

B. Numerical Solution

A direct solution of $(**)$ is possible. Brennan & Schwartz suggest to proceed as follows:

> Solve $Aw = b$ *componentwise* such that
> the side condition $w \geq g$ is obeyed.

This is a somewhat vague outline of an approach: the implementation matters. It is based on the Gaussian elimination, which in its first phase transforms $Aw = b$ into an equivalent system $\tilde{A}w = \tilde{b}$, so that $\tilde{A}$ is a triangular matrix (here bidiagonal). Then, in the second phase, the above principle of Brennan & Schwartz can be solved with one loop. When $\tilde{A}$ is upper triangular, then this loop to *solve* $\tilde{A}w = \tilde{b}$ is a *backward* recursion. For a lower triangle $\tilde{A}$, the loop is *forward*. If in the ith step of the loop $\tilde{w}_i$ denotes the component of the solution of $\tilde{A}w = \tilde{b}$, then $w_i := \max\{\tilde{w}_i, g_i\}$ appears to be a realization.

But w depends on the loop's order. Only one direction works. An implementation must make sure that the characteristic structure of the option is matched.

For a **put** this means:

Let i_f be the index of the node S_i that is closest to the contact point*, where V touches the payoff. Or more definitely, in the transformed variables,

$w_i = g_i$ for $i \le i_\mathrm{f}$, and

$w_i > g_i$ for $i_\mathrm{f} < i \le m$.

This structure is known from the theory, but i_f is unknown. For the put, $w_1 = g_1$ is the starting point, and the $w_i := \max\{\tilde{w}_i,\, g_i\}$-loop is *forward*. Hence $\tilde{A}$ must be a lower triangle, which amounts to an RL-decomposition of A. This establishes the lower triangle $\tilde{A} := L$, and $\tilde{b}$ is solution of $R\tilde{b} = b$. The first components of the loop will be $w_i = g_i$, until the first index with $w_i > g_i$. This fixes the index i_f.

Algorithm (put)

1st phase: Calculate the RL-decomposition of A. Then set $\tilde{A} = L$ and obtain $\tilde{b}$ from $R\tilde{b} = b$ (backward loop).

2nd phase: forward loop, start with $i = 1$. Calculate the next component of $\tilde{A}w = \tilde{b}$; denote it $\tilde{w}_i$.
Set $w_i := \max\{\tilde{w}_i,\, g_i\}$.

---

* provided the chosen interval is large enough, $S_1 < S_\mathrm{f}$

The **costs** are low (solution of a linear system with tridiagonal matrix). It can be shown that the above procedure for a standard option with the underlying matrix A works well.

For a **call** ($\delta > 0$) one proceeds the other way: The loop starts with $w_m = g_m$ and the second phase is a backward loop. To make this possible, in the first phase the "traditional" LR-decomposition of A establishes an upper triangle $\tilde{A} = R$, and $\tilde{b}$ is obtained from $L\tilde{b} = b$ in a foward loop.

Remark on the accuracy:

Since $V(S,t)$ fails to be twice continuously differentiable w.r.t. S at $S_{\mathrm{f}}(t)$, we expect some bad influence on the accuracy. (Recall that Crank–Nicolson even assumes $y \in C^4$.) But in spite of this lack of smoothness, the Crank–Nicolson approach here is sufficiently accurate, oscillations diminish rapidly. The lack of smoothness in the payoff is worse. Even extrapolation works rather well, although the assumptions of smoothness are not satisfied.

Outlook

This concludes the introduction in basic Computational Finance. An essential part of the course are the exercises and the programming assignments.

If this is the material for one semester, then there will be more time left for some additional topics. (In my course, typically, there are two weeks left.) This additional material is not included in these course notes, because it will differ from course to course, depending on the interests and the knowledge of the students. For a textbook explaining additional material, see [Seydel: Tools for Computational Finance, sixth Edition (2017)], from which we take the following section numbers. Possible topics include

- an analytic method: interpolation (§4.8.1), or integral equation (§4.8.4)
- *upwind* scheme and its relevance (§6.4 – 6.5)
- a penalty method in the two-dimensional case (§6.7)
- case studies, such as the two-dimensional tree method of Exercise 6.2
- jump diffusion (§1.9, §7.3)

and for a two-semester course or an accompanying seminar, one may address

- finite elements (Chapter 5)

- nonlinear Black–Scholes problems (§7.1 – 7.2)