Numerical algorithms for R&D stochastic

control models

Yue Kuen KWOK

Department of Mathematics, Hong Kong University of Science and Technology, Hong Kong

* This is a joint work with Chi Man LEUNG.

Overview

- We develop a finite-time real option model with stochastic control that explores the optimal strategy of R&D effort in the development of an innovative product.
- The R&D stochastic control model includes market uncertainty and technological uncertainty, and the firm is allowed to adopt its optimal strategy of R&D effort as control together with the right to abandon the R&D project.
- We model the hazard rate of arrival to be dependent on the current R&D effort and knowledge accumulation in the R&D process, so the hazard rate is non-memoryless.

- We present the HJB formulation of the stochastic control model in combination with the linear complementarity formulation of the optimal stopping rule of abandonment.
- An efficient finite difference algorithm coupled with policy iteration and penalty approximation has been developed to solve for the optimal control strategy of R&D effort.
- Special attention has been taken in the choice of discretization of the HJB equation so that convergence of the numerical solution to the viscosity solution of the HJB equation is guaranteed.
- We performed extensive numerical studies on the optimal control of R&D effort with respect to market conditions and knowledge stock.

R&D stochastic control model

• Stochastic control variable is the rate of expending R&D effort

1. Market uncertainty – stochastic fundamental of profit flow rate

• The profit flow rate x_t generated from the innovative product is

$$dx_t = \mu x_t \ dt + \sigma x_t \ dZ_t,$$

where μ is the constant drift rate, σ is the constant volatility parameter, and Z_t is the standard Brownian motion.

• Reward function

Conditional on $x_t = x$, the expected profit from the product is

$$W(x,t) = E\left[\int_{t}^{T} e^{-r(s-t)} x_{s} \, ds \mid x_{t} = x\right]$$

= $\frac{x}{r-\mu} \left[1 - e^{-(r-\mu)(T-t)}\right], \quad t < T.$

At time t = T, the profit flow terminates forever, so W(x,T) = 0.

Random date of termination (life span) of existing technology

- We model the uncertainty in the life span of the product by assuming the random arrival time T of the termination date to be exponentially distributed with parameter $\lambda = \frac{1}{\overline{T}}$, where \overline{T} is the mean of T.
- The random arrival time T is modeled as the first jump time of a Poisson process with intensity rate λ and this Poisson process is independent of the profit flow process x_t .
- The corresponding reward function W(x) then becomes

$$W(x) = \int_0^\infty \lambda e^{-\lambda u} E\left[\int_0^T e^{-rs} x_s \, ds | T = u\right] du$$
$$= \int_0^\infty \lambda e^{-\lambda u} \left\{\frac{x}{r-\mu} \left[1 - e^{-(r-\mu)u}\right]\right\} du$$
$$= \frac{x}{r+\lambda-\mu}.$$

2. Technological uncertainty

Technological uncertainty is modeled by the *hazard rate* of arrival of discovery of the innovative product.

- Hazard rate increases with the firm's current R&D effort and knowledge stock.
- Let u(t) denote the control variable for the rate of expending R&D effort and z(t) be the path dependent variable of knowledge stock. The hazard rate at time t is modeled by

$$h(t) = au(t) + bz(t), \qquad 0 \le t \le T,$$

where a > 0 and $b \ge 0$ are constant parameters.

3. Knowledge stock and cost function

• The firm's knowledge stock z(t) grows with R&D effort u(t) as

$$\frac{dz}{dt} = u(t), \qquad 0 \le t \le T,$$

where $z(0) = z_0 \ge 0$ is the initial knowledge stock of the firm.

• The rate of cost c(u) is a power function of u.

4. Abandonment right

- We allow the firm to adopt the irreversible decision of abandonment of R&D.
- Even at u = 0, the cost c(0) may remain to be positive (say, maintenance of research facilities), so the abandonment decision helps save the cost of maintaining the R&D process.
- The firm may choose to abandon R&D optimally when x_t falls to a sufficiently low level. The analysis of the abandonment right requires the determination of the optimal stopping rule in the R&D stochastic control problem.

Issues to be addressed

- 1. How would the firm's R&D expenditure evolve under different market conditions, say, in response to the current level and volatility of the profit flow rate, and the remaining life span of the relevant technology?
- 2. How would the firm's R&D expenditure change under varying levels of knowledge stock? Would the firm put off R&D effort when its knowledge stock reaches certain threshold level?

Remark

The model serves the role of examining how different rates of knowledge stock and R&D costs may affect R&D policy decision.

Hamilton-Jacobi-Bellman formulation

Let V(x, z, t) denote the time-t value function of the R&D project conditional on $x_t = x$ and $z_t = z$. Using the Bellman optimality condition, the HJB equation that governs the value function is given by

$$V(x, z, t) = \lim_{dt \to 0} \max(0, \sup_{u \in Q} \{-c(u)dt + h(t)W(x, t)dt + [1 - h(t)dt]e^{-rdt}E\left[V(x_{t+dt}, z_{t+dt}, t+dt)|x_{t} = x, z_{t} = z\right]\})$$

- The optimal stopping rule is applied when the firm either chooses to abandon the project (with zero value being resulted) or continues the R&D process.
- When continuation of the R&D process is optimally chosen, the corresponding optimal control $u^*(t)$ is determined so that the continuation value is maximized.

The continuation value consists of 3 terms:

(i) cost of operating R&D (negative value);

(ii) with probability h(t)dt, R&D succeeds within (t, t + dt) and the expected profit derived from the product is W(x, t);

(iii) with probability 1 - h(t)dt, R&D continues at t + dt and the discounted expected value of the project is given by

$$e^{-rdt}E\left[V(x_{t+dt}, z_{t+dt}, t+dt)|x_t = x, z_t = z\right]$$

By applying Ito's lemma, the last term can be expressed as

$$E\left[V(x_{t+dt}, z_{t+dt}, t+dt)|x_{t} = x, z_{t} = z\right]$$

= $V + \frac{\partial V}{\partial t}dt + \mu x \frac{\partial V}{\partial x}dt + \frac{\sigma^{2}x^{2}}{2} \frac{\partial^{2}V}{\partial x^{2}}dt + u \frac{\partial V}{\partial z}dt + O((dt)^{\frac{3}{2}}).$

The combined HJB formulation of the optimal control on R&D effort and linear complementarity formulation of the optimal stopping rule at abandonment is given by

$$\max(-V, \sup_{u \in Q} \{h(t)W(x,t) - [h(t) + r]V + \frac{\partial V}{\partial t} + \mu x \frac{\partial V}{\partial x} + \frac{\sigma^2}{2} x^2 \frac{\partial^2 V}{\partial x^2} + u \frac{\partial V}{\partial z} - c(u)\}) = 0, \quad x > 0, z \ge 0, t \in [0,T).$$

Auxiliary conditions

• Since the R&D process is sure to terminate at T, we have the terminal condition

$$V(x,z,T)=0$$

for all values of x and z.

• In the stopping region, we have abandonment of the project, so the value function becomes zero for $x \le x^*(t)$. Here, $x^*(t)$ is the stopping boundary at time t.

The far field boundary conditions at $z \to \infty$ and $x \to \infty$ are

(i) At exceedingly high value of z, the hazard rate tends to infinite value. The innovative product is almost surely to be delivered at the next instant, so

$$V(x,z,t)
ightarrow W(x,t)$$
, as $z
ightarrow\infty$.

(ii) At $x \to \infty$, we adopt the linear asymptotic boundary condition on V, where $\frac{\partial^2 V}{\partial x^2} \to 0$. We then have

 $V(x,z,t) \to C_1(z,t)x + C_2(z,t), \qquad \text{as } x \to \infty.$

As $x \to \infty$, one may deduce that the firm increases its R&D effort to the maximum level. The optimal control u^* is independent of x, z and t.

Under these assumptions, we derive the partial differential equations for both coefficient functions $C_1(z,t)$ and $C_2(z,t)$. The closed form analytic formulas are:

$$C_1(z,t) = \frac{1 - e^{-(r-\mu)(T-t)}}{r-\mu} - e^{\frac{(a\hat{u}^* + bz + r-\mu)^2}{2b\hat{u}^*}} \sqrt{\frac{2\pi}{b\hat{u}^*}} [N(d_{11}) - N(d_{12})]$$

and

$$C_2(z,t) = -c(\hat{u}^*)e^{\frac{(a\hat{u}^*+bz+r)^2}{2b\hat{u}^*}}\sqrt{\frac{2\pi}{b\hat{u}^*}}[N(d_{21}) - N(d_{22})],$$

where \hat{u}^* is the supremum value of u within the admissible set of controls, and

$$d_{11} = -\frac{a\hat{u}^* + bz + r - \mu}{\sqrt{b\hat{u}^*}}, \qquad d_{12} = d_{11} - \sqrt{b\hat{u}^*}(T - t),$$
$$d_{21} = -\frac{a\hat{u}^* + bz + r}{\sqrt{b\hat{u}^*}}, \qquad d_{22} = d_{21} - \sqrt{b\hat{u}^*}(T - t).$$

Discretization of the HJB equations using the finite difference approach

- The proposed numerical scheme satisfies the relevant properties of consistency, monotonicity and stability; so the solution to the numerical scheme converges to the viscosity solution of the HJB equations.
- The solution of the non-linear discretized scheme is obtained via the policy iteration method.
- We transform the linear complimentarity formulation into the penalty approximation formulation by appending a penalty term $-\frac{AV}{\varepsilon}$, where $A \in \{0,1\}$ and ε is a sufficiently small parameter. The appended term becomes dominant when the state variables lie in the stopping region in which the R&D process should be optimally abandoned.

The corresponding penalized form can be expressed as

$$F(V) = \sup_{A \in \{0,1\}, u \in Q} \left\{ -\frac{\partial V}{\partial \tau} + \mu x \frac{\partial V}{\partial x} + \frac{\sigma^2}{2} x^2 \frac{\partial^2 V}{\partial x^2} + u \frac{\partial V}{\partial z} - c(u) - [h(\tau) + r]V + h(\tau)W(x,\tau) - \frac{AV}{\varepsilon} \right\} = 0,$$

where $\tau = T - t$.

The corresponding auxiliary conditions are prescribed as follows:

$$V(x, z, 0) = 0, \qquad (x, z) \in (0, \infty) \times [0, \infty),$$

$$V(0, z, \tau) = 0, \qquad (z, \tau) \in [0, \infty) \times [0, T],$$

$$V(x, z, \tau) \rightarrow W(x, \tau) \text{ as } z \rightarrow \infty, \qquad (x, \tau) \in [0, \infty) \times [0, T],$$

$$V(x, z, \tau) = C_1(z, \tau)x + C_2(z, \tau) \text{ as } x \rightarrow \infty, \quad (z, \tau) \in [0, \infty) \times [0, T].$$

Let $V_{j,k}^n$ and W_j^n denote the numerical approximation to $V(x_j, z_k, \tau_n)$ and $W(x_j, \tau_n)$, respectively. We also let $u_{j,k}^n$ and $A_{j,k}^n$ denote the respective control strategy at the nodal point (x_j, z_k, τ_n) .

Fully implicit discretization is adopted and appropriate forward / central / backward differencing is applied to various spatial differential operators so that the condition of positive coefficients is enforced.

The resulting discretized scheme is obtained as follows:

$$\frac{V_{j,k}^{n+1} - V_{j,k}^{n}}{\Delta \tau} = \sup_{\substack{A_{j,k}^{n+1} \in \{0,1\}, u_{j,k}^{n+1} \in Q}} \left\{ a_j V_{j+1,k}^{n+1} - (a_j + b_j + c_{j,k}^{n+1}) V_{j,k}^{n+1} + b_j V_{j-1,k}^{n+1} + d_{j,k}^{n+1} V_{j,k+1}^{n+1} + e_{j,k}^{n+1} \right\},$$

where

$$a_{j} = \frac{\mu x_{j}}{\Delta x} + \frac{\sigma^{2}}{2} \frac{x_{j}^{2}}{\Delta x^{2}}, \quad b_{j} = \frac{\sigma^{2}}{2} \frac{x_{j}^{2}}{\Delta x^{2}},$$
$$c_{j,k}^{n+1} = a u_{j,k}^{n+1} + b z_{k} + r + \frac{u_{j,k}^{n+1}}{\Delta z} + \frac{A_{j,k}^{n+1}}{\varepsilon},$$

$$d_k^{n+1} = \frac{u_{j,k}^{n+1}}{\Delta z}, \quad e_{j,k}^{n+1} = (au_{j,k}^{n+1} + bz_k)W_j^{n+1} - c(u_{j,k}^{n+1}).$$

The coefficients a_j , b_j , $c_{j,k}^{n+1}$ and $d_{j,k}^{n+1}$ are all non-negative. We write the discretized scheme together with the appropriate numerical boundary conditions as

$$G_{j,k}^{n+1}(V_{j,k}^{n+1}, V_{j+1,k}^{n+1}, V_{j-1,k}^{n+1}, V_{j,k+1}^{n+1}, V_{j,k}^{n}) = 0,$$

where $V_{j,k}^{n}$ is the grid value function defined at $(x_{j}, z_{k}, \tau_{n+1})$.

Technical conditions on Δx , Δz and $\Delta \tau$

Recall that $\Delta x = \frac{x_{\max}}{j_{\max}}$, $\Delta z = \frac{z_{\max}}{k_{\max}}$ and $\Delta \tau = \frac{T}{N}$. The stepwidth parameter and time step are chosen such that

$$\Delta x = \beta_1 \delta$$
, $\Delta z = \beta_2 \delta$ and $\Delta \tau = \beta_3 \delta$,

where β_1, β_2 and β_3 are positive constants independent of the small parameter δ .

These assumptions on Δx , Δz and $\Delta \tau$ are necessary in order to establish pointwise consistency of the numerical scheme. In addition, the l_{∞} -stability property and monotonicity property of the numerical scheme can be established.

Proposition The numerical scheme observes the properties of consistency (pointwise), monotonicity and l_{∞} -stability. Provided that the strong comparison property holds, the numerical solution to scheme $G(\cdot)$ converges to the viscosity solution of the HJB formulation F(V) = 0.

Consistency (pointwise), monotonicity and l_{∞} -stablity

1. Consistency

The numerical scheme $G(\cdot)$ is said to be consistent if for any smooth test function $V(x, z, \tau)$ having bounded partial derivatives of all orders, we have

$$\lim_{\delta \to 0} |F(V) - G(\cdot)| = 0.$$

2. Monotonicity

The numerical scheme $G(\cdot)$ is said to be monotone if

 $G_{j,k}^{n+1}(V_{j,k}^{n+1}, \{Y_{a,b}^{n+1}\}_{(a,b)\neq(j,k)}, Y_{j,k}^{n}) \leq G_{j,k}^{n+1}(V_{j,k}^{n+1}, \{X_{a,b}^{n+1}\}_{(a,b)\neq(j,k)}, X_{j,k}^{n})$ where $Y_{j,k}^{n+1} \geq X_{j,k}^{n+1}$, valid for any (j,k).

3. l_{∞} -stability

The numerical scheme $G(\cdot)$ is said to be l_{∞} -stable if $||V^{n+1}||_{\infty} \leq C$ for $0 \leq n \leq N$, $T = N \triangle \tau$, $\Delta \tau \rightarrow 0$, $\Delta x \rightarrow 0$ and $\Delta z \rightarrow 0$, where $||V^{n+1}||_{\infty} = \max_{j,k} \{|V_{j,k}^{n+1}|\}$, and C is independent of Δx , Δz and $\Delta \tau$.

Policy iteration scheme

- We derive the appropriate policy iteration scheme that solves scheme $G(\cdot)$ through an iterative search for $u_{j,k}^{n+1*}$ and $A_{j,k}^{n+1*}$. Noting that the boundary values $V_{j,k_{\max}}^{n+1}$ at $k = k_{\max}$ are known, we proceed to solve for $V_{j,k}^{n+1}$ through marching backward in k, $k = k_{\max} - 1, ..., 1$.
- For a fixed value of k, we solve recursively for the optimal control variables, where each iteration requires numerical solution of a system of $j_{\text{max}} 1$ algebraic equations.

We rewrite scheme $G(\cdot)$ as follows:

$$\sup_{\substack{u_{j,k}^{n+1}, A_{j,k}^{n+1}}} \{a_j \Delta \tau V_{j+1,k}^{n+1} - [1 + (a_j + b_j + c_{j,k}^{n+1}) \Delta \tau] V_{j,k}^{n+1} + b_j \Delta \tau V_{j-1,k}^{n+1} + h_{j,k}^{n+1}\} = 0, \quad j = 1, 2, ..., j_{\text{max}} - 1, \quad (A)$$

where the known quantities are lumped into $h_{j,k}^{n+1}$ defined as

$$h_{j,k}^{n+1} = V_{j,k}^n + \Delta \tau (d_{j,k}^{n+1} V_{j,k+1}^{n+1} + e_{j,k}^{n+1}).$$

In terms of V_k^{n+1} , B_k , h_k^{n+1} , the above scheme (A) can be expressed into the following matrix form:

$$\sup_{\substack{u_{j,k}^{n+1}, A_{j,k}^{n+1}}} \{-B_k \mathbf{V}_k^{n+1} + \mathbf{h}_k^{n+1}\} = \mathbf{0}, \quad k = 0, 1, ..., k_{\max} - 1.$$
(B)

The tridiagonal matrix B_k can be seen to be a *M*-matrix. In the policy iteration scheme presented below, the *M*-matrix property provides a sufficient condition for the convergence of the policy iteration procedure.

Let $(V_k^{n+1})^i$ denote the i^{th} iterate of the vector V_k^{n+1} . The sequence of steps in the policy iteration are outlined as follows:

- 1. Set the initial guess of \mathbf{V}_k^{n+1} to be $(\mathbf{V}_k^{n+1})^0 = \mathbf{V}_k^n$.
- 2. Assuming that the value of $(V_k^{n+1})^i$ is known, the i^{th} iterate of the pair of optimal control variables $(u_{j,k}^{n+1}, A_{j,k}^{n+1})^i$ is determined by

$$(u_{j,k}^{n+1}, A_{j,k}^{n+1})^i = \operatorname*{argmax}_{u \in Q, A \in \{0,1\}} \{ (-B_k (\mathbf{V}_k^{n+1})^i + \mathbf{h}_k^{n+1})_j \},\$$

where $(-B_k(\mathbf{V}_k^{n+1})^i + \mathbf{h}_k^{n+1})_j$ is the j^{th} -component of the vector.

3. Solve the following linear system of equations:

$$-(B_k)^i (\mathbf{V}_k^{n+1})^{i+1} + (\mathbf{h}_k^{n+1})^i = \mathbf{0},$$

where

$$(B_k)^i = B_k|_{(u_{j,k}^{n+1}, A_{j,k}^{n+1})^i}$$
 and $(\mathbf{h}_k^{n+1})^i = \mathbf{h}_k^{n+1}|_{(u_{j,k}^{n+1}, A_{j,k}^{n+1})^i}$.

The policy iteration is terminated when

$$\max_{j} \frac{(V_{j,k}^{n+1})^{i+1} - (V_{j,k}^{n+1})^{i}}{(V_{j,k}^{n+1})^{i+1}} < \text{tolerance value}.$$

Proposition The iterates $(V_k^{n+1})^i$, i = 1, 2, ..., of the iteration algorithm converge to the unique solution of scheme (B) for any initial guess $(V_k^{n+1})^0$.

Numerical tests on convergence of the numerical algorithm

Two major sources of errors: discretization error arising from the discretization of the differential terms and the error arising from approximating the auxiliary conditions via numerical boundary conditions.

The parameter values used in our calculations were chosen to be: r = 0.05, $\mu = 0.01$, $\sigma = 0.3$, a = b = 1, $\varepsilon = 10^{-8}$, T = 1, $x_{\text{max}} = 10$, $z_{\text{max}} = 10$ and $c(u) = 0.01 + \frac{u^2}{2}$. We list the numerical solution values of the value function evaluated at x = 5, z = 5 and T = 1 with varying values of number of time steps.

number of	numerical value	difference in	ratio of
time steps	V(5, 5, 1)	numerical solutions	difference
64	2.1951		
128	2.0060	0.1891	1.9762
256	1.9107	0.0951	1.9843
512	1.8628	0.0479	1.9895
1024	1.8389	0.0239	2.0042

Linear rate of convergence is confirmed since the ratio of difference in numerical solutions is close to 2 when the number of time steps is doubled. The impact of the choices of the upper boundaries, x_{max} and z_{max} , on accuracy of the numerical solutions.

x_{max}	<i>z</i> max	numerical value	difference in
		V(50, 50, 1)	numerical solutions
100	100	48.3371	
200	200	48.3351	0.0020
400	400	48.3338	0.0013
800	800	48.3329	0.0009

When x_{max} and z_{max} are chosen to be about 200, the numerical errors caused by finite truncation of the computational domain is insignificant. We list the numerical solution of V(50, 50, 1) with various choices of x_{max} and z_{max} .

Analysis of the optimal R&D effort

- We explore how the firm *adopts its optimal R&D effort in response to various market conditions* and economic scenarios by performing the sensitivity analysis of the optimal control u^* with respect to different model parameters, like time to expiry, knowledge stock level, volatility of the stochastic state variable, etc.
- We explore the *optimal abandonment policy* adopted by the firm by computing the optimal abandonment boundary that separates the continuation region and abandonment region.
- The following set of parameter values are used in generating the numerical plots: r = 0.05, $\mu = 0.01$, $\sigma = 0.3$, a = b = 1, $\varepsilon = 10^{-8}$, T = 1, $c(u) = 10 + \frac{u^2}{2}$, Q = [0, 10].

Plot of optimal control u^* against time to expiry τ



29

- When $\tau < \tau^*$, the firm optimally chooses to abandon its R&D at any level of the stochastic state variable.
- When τ > τ*, the optimal control u* increases with increasing τ and tends to some asymptotic level when τ is sufficiently large. This is because higher expected profit W(x,τ) is generated from the project as τ increases.
- Since $W(x,\tau)$ is bounded above by $\frac{x}{r-\mu}$ for a fixed value of x, so u^* tends to some asymptotic level as τ increases to some sufficiently high value ($u^* \rightarrow 1.96$ as shown).

Plot of optimal control u^* against current value of the state variable x with varying values of knowledge stock z.



 u^* increases almost linearly with respect to x until up to the level $\hat{u}^* = \sup Q = 10$. The firm optimally increases the R&D effort when z is lower. For a fixed value of x, when z assumes a lower value, the firm increases the control u^* to speed up R&D to increase the expected value of profit.

Plot of optimal control u^* against knowledge stock z with cost function $c(u) = 10 + \frac{u^{1.5}}{2}$.



The optimal control decreases and tends to some asymptotic level with increasing knowledge stock.

Plot of the optimal abandonment threshold x^* against τ with varying values of the fixed cost c_0 in the cost function

$$c(u) = c_0 + \frac{u^2}{2}.$$

Higher fixed cost c_0 , higher optimal abandonment threshold.



- At a given value of τ , the firm optimally chooses to continue the R&D project when the stochastic state variable x_t assumes a value higher than x^* and abandon the project if otherwise.
- In the (τ, x) -plane, the region above (below) the optimal abandonment boundary represents the continuation (abandonment) region.
- The abandonment threshold is seen to be an increasing function of the fixed cost c_0 .
- When the calendar time is sufficiently far from the expiration of the technology, x^* does not depend sensibly on τ .

Summary

- When the calendar time is approaching the expiration date of the technologies, it is optimal for the firm to abandon the R&D project at any level of the profit flow rate. Hence, for finite time horizon R&D models, the right of abandonment is worthy.
- Taking the stochastic profit flow rate as a proxy of the market conditions, the firm increases its R&D effort with an increasing level of profit flow rate and decreasing volatility of the stochastic profit flow rate.
- Our model exhibits a phenomenon similar to the "pure knowledge effect", where the firm may choose optimally to put off R&D effort when the knowledge stock reaches certain threshold value (though the R&D project is kept in progress). This threshold value depends on the current level of the stochastic profit flow rate.

Potential future works – Two-firm real game option model with spillovers of R&D knowledge

- One may consider extending the existing framework of stochastic control R&D model with knowledge stock to R&D race between two competing firms.
- We may include the spillover effects of knowledge accumulation from one firm to its rival. Under a two-firm R&D race model, one has to analyze the various types of strategic equilibriums to be adopted by the two firms and the associated first-mover and second-mover advantages.