# Advertising for a new product introduction: a stochastic approach

Luca Grosset

*[...] we may admit in private that the wind is cold on the peaks of abstraction. The fact that the objects and examples in functional analysis are themselves mathematical theories makes communications with non-mathematicians almost hopeless and deprives us of the feedback that makes mathematics more than an aesthetical game with axioms. The dichotomy between the many small and directly applicable models, and the large, abstract supermodel cannot be explained away. Each must find his own way between Scylla and Charybdis.*

Pedersen, G.K., *Analysis Now*, Springer, New York, 1989.

To my future wife Anna, with love

# PREFACE

Many marketing policies can be correctly explained and analyzed only through a stochastic approach to the problem. In this thesis the planning of a pre-launch publicity campaign has been studied using the stochastic control theory and some recent results of the stochastic linear quadratic control theory. We assume that a firm controls the goodwill evolution of a product through the advertising flow, or some other communication channel, in the programming interval $[0, T]$. The advertising flow increases the goodwill level which otherwise spontaneously decreases. Such hypotheses have been introduced by Nerlove and Arrow in a deterministic framework and have led to the development of a model class called "advertising capital model", in which the advertising flow is considered as an investment in goodwill (a stock which represents the firm image in the market). In this work:

- we assume that the advertising effects on the goodwill evolution are stochastic;
- we describe the effects of the word-of-mouth publicity, by introducing a diffusion term representing the goodwill volatility;
- we assume that the firm can use different advertising channels.

Using the stochastic control theory, under these new assumptions we can introduce a general model that describes the economic problem. The presence of a control in the diffusion term makes the problem mathematically interesting and different from the one studied through the deterministic approach as well as from the one introduced by Tapiero in a stochastic framework. Since the problem is so general, we cannot obtain any solutions in a closed form and it is therefore necessary to specialize it, in order to study some sub-problems which are both simple from a mathematical point of view and relevant for their economic features. At first, we study the introduction of a new product in the market with the aims of maximizing the expected utility coming from the goodwill level at the launch time $T$ and minimizing the total advertising costs. In fact, we summarize these interesting elements for the decision-maker in the same objective functional. The problem is solved using the generalized differential Riccati equation when the utility and the advertising cost functions are quadratic. In this situation the decision-maker is risk proclive and hence the advertising policy is more aggressive than the one obtained in the deterministic case because the introduction of risk in the system is considered as a further utility. Then we consider the possibility of using two different advertising channels for the planning of an event (e.g. a concert or a workshop) characterized by a limited number of seats. For such an event the firm wants to drive the goodwill in order to obtain a final demand as close as possible to the congestion threshold minimizing the total advertising costs. The problem is solved when the advertising cost function and the penalty function are quadratic. The results obtained permit us to compare the efficiency of the two advertising channels and to establish how much the risk connected with the stochastic channel influences its usage. Moreover, we can see that the problem remains well-defined even though the advertising costs of the stochastic channel become negligeable (in such hypotheses the analogous deterministic problem is singular), because the risk connected with this channel can be seen as a further cost. Both problems are solved using some recent results of the stochastic linear quadratic control theory connected with the Riccati equation. This approach permits us to find the optimal control in the feedback form. The comparison between the deterministic and the stochastic results is rather interesting because, when the control affects directly the diffusion term of the motion equation, the two approaches give different optimal policies. This work is composed of two parts. In the first part, which is made up of four Chapters, we introduce the mathematical instruments that are used, while in the second part, which presents the new results, we discuss the marketing models and the solutions obtained. More specifically, Chapter 1 is a summary of the stochastic calculus instruments. In Chapter 2 the general results of the forward and backward stochastic differential equations are presented. The formulation of an optimal control problem in a stochastic framework is the topic of Chapter 3. In the same Chapter we describe the necessary conditions, which consist in Peng's Maximum Principle. In Chapter 4 we deal with the stochastic linear quadratic control problem and we describe the generalized Riccati equation that, under suitable hypotheses, represents necessary and sufficient conditions for the optimality. Chapter 5 opens the second part of this work. Here, we introduce the general model for the launch of a new product in the market. In the following two particular sub-problems are studied: they are interesting for their mathematical characteristics and their economic interpretations. In Chapter 6 we formulate and solve a stochastic extension of the Nerlove and Arrow's advertising model with one communication channel, by assuming that the cost and utility functions are quadratic. In Chapter 7 the communication mix problem for an event planning is introduced and solved. The results obtained permit us to compare the efficiency of different advertising channels. Chapter 6 and 7 are conceived as independent works to permit a separate reading. This choice may give the impression of redundance which is however necessary to make the two Chapters complete and autonomous. Finally, in Chapter 8 we summarize our analysis and present some further research directions which seem promising.

*Acknowledgments*

I wish to thank Alessandra Buratto for all the discussions that led to an improved version of Chapter 7.
I am grateful to Professor Bruno Viscolani for all the suggestions and comments on this work, but, most of all, for his help and his friendship. Without his teaching this work could not have been written.

# PREFAZIONE

Molte politiche di marketing possono essere correttamente spiegate e analizzate solo grazie ad un approccio stocastico al problema. In questo lavoro la pianificazione di una campagna pubblicitaria, che precede l'introduzione di un prodotto nel mercato, è stata studiata utilizzando il controllo stocastico e sfruttando alcuni risultati recenti di controllo stocastico lineare quadratico.

Si suppone che un'azienda controlli, attraverso il flusso di pubblicità o di altre forme di comunicazione, l'evoluzione del goodwill di un prodotto durante l'intervallo di programmazione $[0, T]$. La comunicazione pubblicitaria contribuisce ad incrementare il livello di goodwill (una variabile di stato che riassume gli investimenti pubblicitari dell'azienda) che altrimenti, senza un'azione diretta, è soggetto ad un decadimento spontaneo. Tali ipotesi sono state introdotte da Nerlove e Arrow in un contesto deterministico e hanno portato a sviluppare una famiglia di modelli, chiamati "advertising capital model," nei quali il flusso pubblicitario è considerato come un investimento nello stock goodwill: un titolo che rappresenta l'immagine dell'azienda nel mercato. In questo lavoro:

- si suppone che gli effetti della pubblicità sull'evoluzione del goodwill siano stocastici;
- si descrive l'effetto della pubblicità "passaparola" con l'inserimento di un termine di diffusione che rappresenta la volatilità del goodwill;
- si suppone che l'azienda possa operare con diversi canali pubblicitari sull'evoluzione del goodwill.

Queste novità portano alla formulazione di un modello generale che descrive il problema economico utilizzando la teoria del controllo stocastico: la presenza del controllo nel termine di diffusione rende il problema matematicamente interessante e sostanzialmente diverso sia da quello affrontato nell'approccio deterministico, sia da quello proposto in termini stocastici da Tapiero. La generalità del modello non permette di ottenere soluzioni in forma chiusa ed è quindi necessario specializzarlo trattando dei casi particolarmente semplici, ma che mantengono un'interpretazione economica interessante.

Inizialmente si studia l'introduzione nel mercato di un nuovo prodotto con l'obiettivo congiunto di massimizzare l'utilità attesa, derivante dal valore del goodwill all'istante di lancio $T$, e di minimizzare la spesa totale in comunicazione. In realtà si includono entrambi questi elementi di interesse per il decisore in un unico funzionale obiettivo. Il problema viene risolto nel caso di costi ed utilità quadratici, utilizzando una generalizzazione dell'equazione di Riccati. In questa situazione il decisore è propenso al rischio e quindi la politica pubblicitaria risulta più aggressiva rispetto a quella che si ottiene nel caso puramente deterministico poiché l'introduzione di elementi di rischio nel sistema viene vista come una possibile opportunità di utilità aggiuntiva.

In seguito si prende in considerazione il possibile utilizzo di due forme di comunicazione per l'organizzazione di un evento (e.g. un concerto oppure un workshop), caratterizzato da un numero limitato di posti disponibili, per il quale si vuole che la domanda sia il più vicino possibile alla soglia di congestione, ancora compatibilmente con la spesa totale in comunicazione. Il problema viene risolto nel caso di funzione costo e funzione di penalità quadratiche: i risultati ottenuti permettono di confrontare l'efficacia dei due canali pubblicitari e di stabilire quanto il rischio connesso con il canale stocastico influenzi il suo utilizzo. Inoltre si vede come, anche supponendo che il costo del canale stocastico diventi nullo, il problema continui ad essere ben posto (in tali ipotesi l'analogo problema deterministico diventa singolare) poiché il rischio connesso con l'azione del decisore può essere interpretato come un costo aggiuntivo.

Entrambi i problemi sono affrontati utilizzando alcuni risultati recenti della teoria del controllo stocastico lineare quadratico e della generalizzazione dell'equazione di Riccati. Questo approccio permette di calcolare il controllo ottimo e di esprimerlo in forma di feedback. Particolarmente interessante risulta il confronto fra le soluzioni ottenute in ambiente stocastico e quelle ricavate supponendo che la legge di evoluzione del goodwill sia l'analoga deterministica. Infatti si osservano notevoli differenze tra i due approcci quando il parametro di controllo influisce direttamente sul termine di diffusione dell'equazione del moto.

Questo lavoro è diviso in due parti: nella prima, che comprende i primi quattro capitoli, si introducono gli strumenti matematici utilizzati, nella seconda, che presenta i contributi originali, vengono discussi i modelli nell'ambito del marketing e i risultati ottenuti. Scendendo più in dettaglio il Capitolo 1 rappresenta un riassunto degli strumenti di calcolo stocastico utilizzati. Nel Capitolo 2 vengono presentati i risultati generali legati alla teoria delle equazioni differenziali stocastiche sia forward che backward. La formulazione in ambito stocastico di un problema di controllo ottimo è l'argomento del Capitolo 3, dove vengono anche descritte le condizioni necessarie per l'ottimalità che consistono nel Principio del Massimo dimostrato da Peng. Nel Capitolo 4 ci occupiamo di problemi di controllo lineare quadratico e descriviamo la generalizzazione dell'equazione di Riccati che, sotto le opportune ipotesi, rappresenta condizioni sia necessarie che sufficienti per l'ottimalità. Nel Capitolo 5, che apre la seconda parte, viene introdotto il modello generale per il lancio di un nuovo prodotto nel mercato, mentre nei due capitoli seguenti vengono trattati due sottoproblemi che ammettono una buona trattabilità matematica e che mantengono una interessante interpretazione economica. Nel Capitolo 6 è formulata e risolta un'estensione stocastica del modello di Nerlove e Arrow con un solo canale pubblicitario, supponendo che le funzioni di costo e di utilità siano quadratiche. Nel Capitolo 7 viene presentato e risolto il problema di marketing mix legato alla campagna di preparazione di un evento. I risultati ottenuti ci permettono di confrontare l'efficacia di diverse forme di comunicazione. I Capitoli 6 e 7 sono stati concepiti come indipendenti per permettere una lettura parziale di questo lavoro. Questa scelta può creare ridondanza, ma è necessaria per rendere i due capitoli completi ed autonomi. Infine, nel Capitolo 8, sono riassunti i risultati dell'analisi effettuata e sono proposte alcune direzioni di ricerca che appaiono particolarmente interessanti.

# Contents

CHAPTER 1

# Stochastic Calculus

## 1. Notation

Given a probability space $(\Omega, \mathcal{E}, \mathbb{P})$ and a time interval $[0, T]$, a stochastic process is a function

$$X_t(\omega) : [0, T] \times \Omega \to \mathbb{R}^n \ .$$

For every fixed $\bar{\omega} \in \Omega$ the function $X_t(\bar{\omega}) : [0, T] \to \mathbb{R}^n$ is a sample path or a trajectory of the process. Generally, the probability space $(\Omega, \mathcal{E}, \mathbb{P})$ is equipped with a filtration $\{\mathcal{G}_t\}_{t \in [0,T]}$, which is an increasing sequence of sub-$\sigma$-fields of $\mathcal{E}$: for all $t, s \in [0, T]$ with $t < s$

$$\begin{aligned} &\mathcal{G}_t \subset \mathcal{E} \ , \\ &\mathcal{G}_t \subset \mathcal{G}_s \ . \end{aligned}$$

The filtration models the arrival of the information during the time interval $[0, T]$: at the time $\bar{t}$ we exactly know whether an event $G \in \mathcal{G}_{\bar{t}}$ occurs or does not occur. Usually, the filtration is required to satisfy some technical assumptions:

**completeness:** $\mathcal{G}_0$ contains all the $\mathbb{P}$-null sets of $\mathcal{E}$;

**right continuity:** for every fixed $\bar{t} \in [0, T]$ and for every $\varepsilon \in (0, T - \bar{t})$ the following relation holds:

$$\bigcap_{t \in (\bar{t}, \bar{t}+\varepsilon)} \mathcal{G}_t = \mathcal{G}_{\bar{t}} \ .$$

A filtration that satisfies these conditions (called "usual conditions") is called standard, and a probability space equipped with a standard filtration is called a standard filtered probability space.

A stochastic process is adapted to a filtration if and only if for every fixed $\bar{t} \in [0, T]$ the random variable $X_{\bar{t}}(\omega)$ is measurable with respect to the $\sigma$-field $\mathcal{G}_{\bar{t}}$. If a process is adapted, then for every fixed $\bar{t} \in [0, T]$ we can use the information contained in $\mathcal{G}_{\bar{t}}$, and therefore we know the true path of the process up to the time $\bar{t}$.

In the following, we always work in a filtered probability space

$$\left(\Omega, \mathcal{E}, \{\mathcal{G}_t\}_{t \in [0,T]}, \mathbb{P}\right) \ ,$$

where a Wiener process $W_t(\omega)$ is defined: i.e. an adapted stochastic process that has the following characteristics:

- $\mathbb{P}(\{\omega \in \Omega : W_0(\omega) = 0\}) = 1$;
- for all $s, t \in [0, T]$, $s \leq t$, the random variable $W_t(\omega) - W_s(\omega)$ is independent of $\mathcal{G}_s$;
- for all $s, t \in [0, T]$, $s \leq t$, the random variable $W_t(\omega) - W_s(\omega)$ is normally distributed with mean 0 and covariance matrix $(t-s)\mathbb{I}$ ($\mathbb{I}$ is the identity matrix).

In other words, the third property means that for all $s, t \in [0, T]$, if $s \leq t$ then

$$\mathbb{E}\left(W_t(\omega) - W_s(\omega) \middle| \mathcal{G}_s\right) = 0 \ ,$$
$$\mathbb{E}\left(\left(W_t(\omega) - W_s(\omega)\right)\left(W_t(\omega) - W_s(\omega)\right)' \middle| \mathcal{G}_s\right) = (t-s)\mathbb{I} \ .$$

It can be proved [**6**, p.25], [**28**, p.21] that a Wiener process $W_t(\omega)$ has almost surely continuous trajectories: i.e. there exists $N \in \mathcal{E}$ such that $\mathbb{P}(N) = 0$ and for all $\bar{\omega} \in \Omega \setminus N$ the function $W_t(\bar{\omega}) : [0, T] \times \Omega \to \mathbb{R}^n$ is continuous. Moreover, given a $m$-dimensional Wiener process $W_t(\omega) = \left(W_t^1(\omega), ..., W_t^m(\omega)\right)$, we can show that each component is an independent one-dimensional Wiener process with respect to the same filtration [**14**, p.6].

The filtration generated by the process $W_t(\omega)$ is defined as follows:

$$\mathcal{F}_t^W = \sigma\{W_s(\omega) : 0 \leq s \leq t\}$$

(it is the smallest sub-$\sigma$-field of $\mathcal{E}$ which makes all the random variables $W_s(\omega)$ measurable for all $s \in [0, t]$). In order to obtain a standard filtration we have to augment the filtration $\mathcal{F}_t^W$ with the set $\mathcal{N}_\mathcal{E}$ which contains all the $\mathbb{P}$-null sets of the $\sigma$-field $\mathcal{E}$:

$$\mathcal{F}_t = \sigma\left\{\mathcal{F}_t^W \cup \mathcal{N}_\mathcal{E}\right\} \ .$$

It can be proved that the augmented natural filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ is standard and that the process $W_t(\omega)$ is a Wiener process even with respect to the augmented natural filtration [**6**, p.58], [**28**, p.23].

Remark 1. *In the following, we always implicitly assume that we work in a probability space $(\Omega, \mathcal{E}, \mathbb{P})$ where a Wiener process $W_t(\omega)$ is defined and where the chosen filtration is the augmented natural one.*

A spontaneous question is: “When do two stochastic processes model the same phenomenon?” Several possible definitions are now introduced.

Definition 1. *Two stochastic processes $X_t(\omega), Y_t(\omega)$ are stochastically equivalent if and only if for all fixed $\bar{t} \in [0, T]$ there exists an* *event* $N_{\bar{t}} \in \mathcal{E}$ *such that* $\mathbb{P}(N_{\bar{t}}) = 0$ *and for all* $\omega \in \Omega \setminus N_{\bar{t}}$ *the following equation holds:*

$$X_{\bar{t}}(\omega) = Y_{\bar{t}}(\omega) \ .$$

*Moreover, two stochastic processes* $X_t(\omega), Y_t(\omega)$ *are indistinguishable if and only if there exists an event* $N \in \mathcal{E}$ *such that* $\mathbb{P}(N) = 0$ *and for all* $\omega \in \Omega \setminus N$ *and for all* $t \in [0, T]$ *the following equation holds:*

$$X_t(\omega) = Y_t(\omega) \ .$$

Clearly, if two stochastic processes are indistinguishable, then they are stochastically equivalent. The converse is not necessarily true, because the more-than-countable union of events with 0 probability may not be an event of probability zero. These two definitions are equivalent when a regularity condition for the trajectories is assumed.

Definition 2. *A stochastic process* $X_t(\omega)$ *is said to be continuous if and only if there exists an event* $N \in \mathcal{E}$ *such that* $\mathbb{P}(N) = 0$ *and for all fixed* $\bar{\omega} \in \Omega \setminus N$ *the function* $X_t(\bar{\omega}) : [0, T] \to \mathbb{R}$ *is continuous.*

Proposition 1. *Let* $X_t(\omega), Y_t(\omega)$ *be two stochastic processes defined on the same filtered probability space. Let us assume that these two processes are continuous and stochastically equivalent, then they are indistinguishable, too.*

*Proof.* [**6**, p.29], [**11**, p.2]. □

We work with the space of the adapted processes with continuous trajectories, and we identify two indistinguishable processes. The following position clarifies this matter: let us consider the space of all continuous and adapted processes on the same filtered probability space. The stochastic equivalence of processes is an equivalence relation, therefore we can consider the quotient space of all continuous and adapted processes by the stochastic equivalence relation. We use the symbol $\mathcal{C}$ to denote this quotient space. Clearly, the Wiener process $W_t(\omega)$ is an element of the set $\mathcal{C}$.
We want to model the evolution of a system and we require the system state functions to be elements of the space $\mathcal{C}$. On the other hand, we need to define some other process space in order to represent the control functions, through which the system can be driven.

Definition 3. *A stochastic process* $X_t(\omega)$ *is said to be measurable if and only if the map*

$$X_t(\omega) : [0, T] \times \Omega \to \mathbb{R}^n$$

*defined on the measurable space* $([0,T] \times \Omega, \mathcal{B}([0,T]) \otimes \mathcal{E})$ *is measurable (by the symbol* $\mathcal{B}([0,T])$ *we denote the Borel* $\sigma$*-field defined on the interval* $[0,T]$*).*
*A stochastic process* $X_t(\omega)$ *is said to be progressively measurable if and only if for all fixed* $\bar{t} \in [0,T]$ *the map*

$$X_t(\omega) : [0,\bar{t}] \times \Omega \to \mathbb{R}^n$$

*defined on the measurable space* $([0,\bar{t}] \times \Omega, \mathcal{B}([0,\bar{t}]) \otimes \mathcal{F}_{\bar{t}})$ *is measurable.*

If we want to define the stochastic integral, it is fundamental to work with processes which are progressively measurable. Fortunately, the processes we are working on are sufficiently regular for this condition to be always satisfied.

PROPOSITION 2. *If* $X_t(\omega) \in \mathcal{C}$ *then* $X_t(\omega)$ *is progressively measurable (every element of an equivalence class of the space* $\mathcal{C}$ *is progressively measurable).*

*Proof.* [**6**, p.23], [**28**, p.17]. □

Let us define some new useful process spaces. First of all, we have to define a new equivalence relation between processes.

DEFINITION 4. *Let* $X_t(\omega), Y_t(\omega)$ *be two measurable and adapted processes defined on the same standard filtered probability space; they are almost everywhere identical if and only if there exists an element* $N \in \mathcal{B}([0,T]) \otimes \mathcal{E}$ *such that* $(\lambda \otimes \mathbb{P})(N) = 0$ *and for all* $(t,\omega) \in ([0,T] \times \Omega) \setminus N$

$$X_t(\omega) = Y_t(\omega) \ .$$

Let us consider the space of all measurable and adapted processes defined on the same standard filtered probability space

$$\left(\Omega, \mathcal{E}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right) .$$

We decide to identify two processes when they are almost everywhere identical (we are considering the quotient space with respect to the almost everywhere identity). We define $\mathcal{L}^p$ (with $p = 1, 2$) as the set of all the equivalence classes $X_t(\omega)$ such that

$$\int_0^T \|X_t(\omega)\|^p \, dt < +\infty \ , \quad \mathbb{P}\text{-almost surely.}$$

$X_t(\omega)$ may be an $n \times m$-dimensional matrix process: in that case, we consider the matrix norm ([**28**, p.354])

$$\|X_t(\omega)\| = \operatorname{tr}\left(X_t(\omega) X_t(\omega)'\right) \ ,$$

and we note that $X_t(\omega) \in \mathcal{L}^p$ if and only if each component process $X_t^{i,j}(\omega) \in \mathcal{L}^2$ (for all $i = 1, ..., n$, and for all $j = 1, ..., m$).
It is interesting to note that if $X_t(\omega) \in \mathcal{L}^1$ and $Y_t(\omega) \in \mathcal{C}$, then $X_t(\omega)\ \ Y_t(\omega) \in \mathcal{L}^1$. If we consider a measurable and adapted process $X_t(\omega)$, it may not be progressively measurable, but there exists another process $Y_t(\omega)$ stochastically equivalent to $X_t(\omega)$ which is progressively measurable [**14**, p.45]. It follows that as long as we identify processes that are almost everywhere identical, there is no difference between adapted measurable processes and progressively measurable processes: in every equivalence class of $\mathcal{L}^1$ there exists an element which is not only adapted and measurable, but also progressively measurable. Therefore, we assume that, when we choose an element of an equivalent class in $\mathcal{L}^1$, it is always a progressively measurable process.

For the processes which are in $\mathcal{L}^1$ we can define the time integral processes by means of the pathwise time integral.

PROPOSITION 3. *There exists a linear function of the linear space $\mathcal{L}^1$ into $\mathcal{C}$, defined as follows:*

$$\begin{array}{ccc} \mathcal{L}^1 & \rightarrow & \mathcal{C} \\ X_t(\omega) & \mapsto & \left(t \mapsto \int_0^t X_s(\omega)\,ds\right) \end{array} .$$

*Proof.* The function is well-defined because: given a progressively measurable element $X_t(\omega)$ of an equivalence class of $\mathcal{L}^1$, the time integral $t \mapsto \int_0^t X_s(\omega)\,ds$ is a well-defined, continuous and adapted process [**14**, p.31]. Moreover, two processes have indistinguishable time integrals if and only if they are almost everywhere identical [**14**, p.31]. Finally, the function is linear because of the linearity of the Lebesgue integral. □

Now, let $n = 1$, so that we consider one-dimensional processes. First of all we define the stochastic integral for such processes, then we extend the definition to the multidimensional processes. Let $k$ be a natural number, and let $t_0, t_1, ..., t_k$ be elements of the set $[0, T]$ such that $0 = t_0 < t_1 < ... < t_k = T$; we define the following function as a one-dimensional simple process:

$$X_t(\omega) = \sum_{i=0}^{k-1} Y_i(\omega)\, I_{(t_i, t_{i+1}]}(t) \ ,$$

where $Y_i(\omega) \in L^2(\Omega, \mathcal{F}_i, \mathbb{P})$ and $I_{(t_i,t_{i+1}]}(t)$ is the indicator function of the interval $(t_i, t_{i+1}]$. First of all, we define the stochastic integral on $[0, \bar{t}]$ ($\bar{t}$ is a fixed time in the interval $[0, T]$) for a simple process. It is a random variable defined $\mathbb{P}$-almost surely by the relation

$$\begin{aligned}&\int_0^{\bar{t}} \left(\sum_{i=0}^{k-1} Y_i(\omega)\, I_{(t_i,t_{i+1}]}(s)\right) dW_s(\omega)\\ &\triangleq \sum_{i=0}^{k-1} Y_i(\omega)\left(W_{\bar{t}\wedge t_{i+1}}(\omega) - W_{\bar{t}\wedge t_i}(\omega)\right) .\end{aligned}$$

Now, let $X_t(\omega)$ be a one-dimensional measurable and adapted process such that

$$\int_0^T |X_t(\omega)|^2\, dt < +\infty \ , \quad \mathbb{P}\text{-almost surely} . \tag{1}$$

May we define the stochastic integral of $X_t(\omega)$ on the interval $[0,\bar{t}]$ using the definition just given for the simple process? The following theorem provides an answer.

THEOREM 4. *Let $X_t(\omega)$ be a measurable and adapted process that satisfies (1), then*

- *there exists a sequence $\left\{X_t^{(h)}(\omega)\right\}_{h\in\mathbb{N}}$ of simple processes such that, as $h\to+\infty$,*

$$\int_0^{\bar{t}} \left|X_s^{(h)}(\omega) - X_s(\omega)\right|^2 ds \quad \xrightarrow{\mathbb{P}} \quad 0 \tag{2}$$

  *(i.e. there exists an approximation with simple processes of $X_t(\omega)$);*
- *for every sequence $\left\{X_t^{(h)}(\omega)\right\}_{h\in\mathbb{N}}$ of simple processes as in (2) there exists one and only one random variable $I(\omega)$ (defined $\mathbb{P}$-almost surely) such that, as $h\to+\infty$,*

$$\int_0^{\bar{t}} X_s^{(h)}(\omega)\, dW_s(\omega) \quad \xrightarrow{\mathbb{P}} \quad I(\omega) \triangleq \int_0^{\bar{t}} X_s(\omega)\, dW_s(\omega) \tag{3}$$

  *(i.e. there exists a natural candidate to be the stochastic integral of the process $X_t(\omega)$);*
- *if two different sequences $\left\{X_t^{(h)}(\omega)\right\}_{h\in\mathbb{N}}$ and $\left\{Y_t^{(h)}(\omega)\right\}_{h\in\mathbb{N}}$ approximate the same process $X_t(\omega)$, in the sense given by the relation (2), then the sequences of their stochastic integrals converge to the same random variable as $h\to+\infty$:*

$$\left.\begin{array}{ll}\int_0^{\bar{t}} X_s^{(h)}(\omega)\, dW_s(\omega) & \xrightarrow{\mathbb{P}} \\ \int_0^{\bar{t}} Y_s^{(h)}(\omega)\, dW_s(\omega) & \xrightarrow{\mathbb{P}}\end{array}\right\} \quad I(\omega) \triangleq \int_0^{\bar{t}} X_t(\omega)\, dW_s(\omega)$$

  *(i.e. the random variable $I(\omega)$ found in (3) does not depend on the choice of the sequence $\left\{X_t^{(h)}(\omega)\right\}_{h\in\mathbb{N}}$).*

*Proof.* [**14**, p.36] □

Now, we have the definition of stochastic processes for one-dimensional processes. What happens if $X_t(\omega)$ is an $n \times m$-dimensional process or $W_t(\omega)$ is an $m$-dimensional Wiener process? First of all, we recall that $X_t(\omega) \in \mathcal{L}^2$ if and only if each component process $X_t^{i,j}(\omega) \in \mathcal{L}^2$ (for all $i = 1, ..., n$, and for all $j = 1, ..., m$). Therefore, we define

$$\int_0^{\bar{t}} X_s(\omega)\, dW_s(\omega) \triangleq \begin{pmatrix} \sum_{j=1}^m \int_0^{\bar{t}} X_s^{1,j}(\omega)\, dW_s^j(\omega) \\ \cdots \\ \sum_{j=1}^m \int_0^{\bar{t}} X_s^{n,j}(\omega)\, dW_s^j(\omega) \end{pmatrix} .$$

We note that in this definition we use the one-dimensional stochastic integral introduced above.
Now, we want to consider the stochastic integral as an $n$-dimensional process depending on the extreme of integration $\bar{t}$.

PROPOSITION 5. *Let $X_t(\omega), Y_t(\omega) \in \mathcal{L}^2$, then the following properties hold*

**linearity:** *for all $a, b \in \mathbb{R}$*

$$\begin{aligned} &\textstyle\int_0^{\bar{t}} aX_s(\omega) + bY_s(\omega)\, dW_s(\omega) \\ &= a \textstyle\int_0^{\bar{t}} X_s(\omega)\, dW_s(\omega) + b \int_0^{\bar{t}} Y_s(\omega)\, dW_s(\omega) \ \ ; \end{aligned}$$

**time consistency:** *for all fixed $\bar{t} \in [0, T]$*

$$\int_0^{\bar{t}} X_s(\omega)\, dW_s(\omega) = \int_0^T I_{[0,\bar{t}]}(s)\, X_s(\omega)\, dW_s(\omega) \ \ ;$$

**adaptivity:** *the process $(\omega, t) \mapsto \int_0^t X_s(\omega)\ dW_s(\omega)$ is adapted to the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$;*

**continuity:** *the stochastic process $(\omega, t) \mapsto \int_0^t X_s(\omega)\, dW_s(\omega)$ is continuous and identifies one and only one element of the space $\mathcal{C}$.*

*Proof.* [**14**, p.37-38] □

If we choose two measurable and adapted processes almost everywhere identical, which are in $\mathcal{L}^2$, then their stochastic integral processes are linked together by the following result.

PROPOSITION 6. *The stochastic integral processes of two measurable and adapted processes, which are in $\mathcal{L}^2$, are indistinguishable if and only if $X_t(\omega), Y_t(\omega)$ are almost everywhere identical.*

*Proof.* [**14**, p.38] □

Therefore, the following linear function is well-defined:

$$\begin{array}{ccc} \mathcal{L}^2 & \to & \mathcal{C} \\ X_t(\omega) & \mapsto & \left(t \mapsto \int_0^t X_s(\omega)\, dW_s(\omega)\right) \end{array} .$$

Now, let us consider a subspace of the process space $\mathcal{L}^2$ defined as follows:

$$\mathcal{H}^2 \triangleq \left\{ X_t(\omega) \in \mathcal{L}^2 : \mathbb{E}\left( \int_0^T \|X_t(\omega)\|^2 \, dt \right) < +\infty \right\} .$$

By the definition given for the matrix norm of matrix processes, we have that $X_t(\omega) \in \mathcal{H}^2$ if and only if each component process $X_t^{i,j}(\omega) \in \mathcal{H}^2$ (for all $i = 1, ..., n$, and for all $j = 1, ..., m$). Clearly, $\mathcal{H}^2$ is a proper subset of $\mathcal{L}^2$ because the time integral $\int_0^T \|X_t(\omega)\|^2 \ dt$ may take finite values with probability one and yet have infinite expectation. As $\mathcal{H}^2 \subset \mathcal{L}^2$, the stochastic integral is well defined also on the elements of $\mathcal{H}^2$. Furthermore, the stochastic integral process defined on an element of $\mathcal{H}^2$ has some special properties.

Proposition 7. *Let $X_t(\omega), Y_t(\omega)$ be elements (with the same dimension) of the process space $\mathcal{H}^2$, then*

(1) *the process $\int_0^t X_s(\omega)\, dW_s(\omega) \in \mathcal{C}$ is a martingale with respect to the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$: i.e. for all $t, s \in [0,T]$ with $s < t$*

$$\mathbb{E}\left( \int_0^t X_u(\omega) \ dW_u(\omega) \ \mid \mathcal{F}_s \right) = \int_0^s X_u(\omega)\, dW_u(\omega)$$

*in particular the mean of the stochastic integral is always zero;*

(2) *the process $\left(\int_0^t X_s(\omega)\, dW_s(\omega)\right)\left(\int_0^t Y_s(\omega)\, dW_s(\omega)\right)' \in \mathcal{C}$ is integrable and the following relation (called Itô isometry) holds for all $s, t \in [0,T]$ with $s \le t$*

$$\begin{aligned} &\mathbb{E}\left[ \left(\int_0^t X_u(\omega)\, dW_u(\omega)\right)\left(\int_0^t Y_u(\omega)\, dW_u(\omega)\right)' \middle| \mathcal{F}_s \right] \\ &= \int_s^t \mathbb{E}\left[ X_u(\omega)\, Y_u(\omega)' \middle| \mathcal{F}_s \right] du \ . \end{aligned}$$

*Proof.* [**14**, p.43] □

The space $\mathcal{H}^2$ is important not only for the previous relation, but also for the following theorem. Let $\mathcal{M}_c^2$ be the space of processes belonging to $\mathcal{C}$ which are square integrable martingale w.r.t. $\{\mathcal{F}_t\}_{t\in[0,T]}$, i.e. the process $M_t(\omega) \in \mathcal{C}$ such that for all $s, t \in [0,T]$, $s < t$,

$$\begin{aligned} &\mathbb{E}\left( M_t(\omega) | \mathcal{F}_s\right) = M_s(\omega) \ , \\ &\mathbb{E}\left(\int_0^T \|M_t(\omega)\|^2 dt\right) < +\infty \ . \end{aligned}$$

We note that $\mathcal{M}_c^2$ is a subspace of the space $\mathcal{C}$. A natural question is: "When does a stochastic integral operator give a process belonging to $\mathcal{M}_c^2$?" The following result answers this question and characterize the image of the space $\mathcal{H}^2$ with respect to the action of the stochastic integral operator.

Theorem 8. *Let $M_t(\omega)$ be an element of the process space $\mathcal{M}_c^2$ , then there exists one and only one element $X_t(\omega)$ of $\mathcal{H}^2$ such that:*

$$M_t(\omega) = m + \int_0^t X_s(\omega)\, dW_s(\omega) \ .$$

*Proof.* [**14**, p.44] □

This result is called martingale representation theorem and it can be explained saying that the image of the space $\mathcal{H}^2$ with respect to the action of the stochastic integral operator is the space of the continuous square integrable martingales which vanish at the time 0.

## 2. Stochastic Calculus

Definition 5. *A process $X_t(\omega) \in \mathcal{C}$ is called an Itô process if and only if there exist two processes $X_t^{\text{drift}}(\omega) \in \mathcal{L}^1$ and $X_t^{\text{diff}}(\omega) \in \mathcal{L}^2$ such that*

$$X_t(\omega) = \bar{x} + \int_0^t X_s^{\text{drift}}(\omega)\, ds + \int_0^t X_s^{\text{diff}}(\omega)\, dW_s(\omega) \ .$$

Clearly, the set of Itô processes is a subspace of the process space $\mathcal{C}$. If $X_t(\omega) \in \mathcal{C}$ is an Itô process, then we say that its stochastic differential is

$$dX_t(\omega) = X_t^{\text{drift}}(\omega)\, dt + X_t^{\text{diff}}(\omega)\, dW_t(\omega) \ .$$

This definition is well-posed because the following proposition holds.

Proposition 9. *Let $X_t(\omega), Y_t(\omega) \in \mathcal{C}$ be two Itô processes:*

$$\begin{aligned} X_t(\omega) &= \bar{x} + \int_0^t X_s^{\text{drift}}(\omega)\, ds + \int_0^t X_s^{\text{diff}}(\omega)\, dW_s(\omega) \ , \\ Y_t(\omega) &= \bar{y} + \int_0^t Y_s^{\text{drift}}(\omega)\, ds + \int_0^t Y_s^{\text{diff}}(\omega)\, dW_s(\omega) \end{aligned}$$

*if they are indistinguishable, then $\bar{x} = \bar{y}$ in $\mathbb{R}$, $X_t^{\text{drift}}(\omega) = Y_t^{\text{drift}}(\omega)$ in $\mathcal{L}^1$, and $X_t^{\text{drift}}(\omega) = Y_t^{\text{drift}}(\omega)$ in $\mathcal{L}^2$.*

*Proof.* [**14**, p.60] □

This result assures the uniqueness of the stochastic differential. The power of the stochastic calculus is connected with the existence of some differential rules which are very similar to the deterministic ones.

PROPOSITION 10. *Let* $X_t(\omega), Y_t(\omega) \in \mathcal{C}$ *be two Itô processes such that*

$$\begin{array}{l} X_t(\omega) = \bar{x} + \int_0^t X_s^{\text{drift}}(\omega)\, ds + \int_0^t X_s^{\text{diff}}(\omega)\, dW_s(\omega) \ \ , \\ Y_t(\omega) = \bar{y} + \int_0^t Y_s^{\text{drift}}(\omega)\, ds + \int_0^t Y_s^{\text{diff}}(\omega)\, dW_s(\omega) \ \ , \end{array}$$

*then the following properties hold:*

**linearity:** *for all* $a, b \in \mathbb{R}$

$$d\left[aX_t(\omega) + bY_t(\omega)\right] = a dX_t(\omega) + b dY_t(\omega) \ \ ;$$

**inner product:** *the stochastic differential of the inner product process* $\langle X_t(\omega), Y_t(\omega)\rangle$ *can be computed as follows (*$X_t(\omega)$ *and* $Y_t(\omega)$ *are* $n$*-dimensional processes, but* $W_t(\omega)$ *is a one-dimensional process):*

$$\begin{array}{l} d\langle X_t(\omega), Y_t(\omega)\rangle = \langle Y_t(\omega), dX_t(\omega)\rangle \\ \qquad + \langle X_t(\omega), dY_t(\omega)\rangle + \left\langle X_t^{\text{diff}}(\omega), Y_t^{\text{diff}}(\omega)\right\rangle dt \ ; \end{array}$$

**chain rule:** *let* $g(x) \in C^2(A, \mathbb{R})$ *be a real function (*$A$ *is an open set of* $\mathbb{R}^n$*); if* $\mathbb{P}\{\omega \in \Omega : \forall t \in [0, T] \ \ X_t(\omega) \in A\} = 1$*, then*

$$\begin{array}{l} d\left[g(X_t(\omega))\right] = \\ \qquad \nabla g(X_t(\omega))\, X_t^{\text{drift}}(\omega)\, dt + \\ \qquad + \frac{1}{2}\mathrm{tr}\left[X_t^{\text{diff}}(\omega)' \,\mathrm{Hess}\left[g(X_t(\omega))\right] X_t^{\text{diff}}(\omega)\right] dt + \\ \qquad + \nabla g(X_t(\omega))\, X_s^{\text{diff}}(\omega)\, dW_t(\omega) \ \ . \end{array}$$

*Proof.* [**6**, p.131], [**14**, p.54], [**28**, p.37]. □

CHAPTER 2

# Stochastic Differential Equations

## 1. Stochastic Differential Equations

Given a $\left(\Omega, \mathcal{E}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ a stochastic differential equation is an object defined as follows:

$$\begin{cases} dX_t(\omega) = b(t,\omega,X_t(\omega))\,dt + \sigma(t,\omega,X_t(\omega))\,dW_t(\omega) \ , \\ X_t(\omega) = \bar{x} \ , \end{cases} \tag{4}$$

where $b : [0,T] \times \Omega \times \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : [0,T] \times \Omega \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$, while $W_t(\omega)$ is an $m$-dimensional Wiener process. As in the deterministic case, the meaning of (4) becomes clear when we define the concept of solution for a stochastic differential equation (SDE for short).

Definition 6. *A process $X_t(\omega) \in \mathcal{C}$ is a solution of the SDE (4) if and only if*

$$X_t(\omega) = \bar{x} + \int_0^t b(s,\omega,X_s(\omega))\,ds + \int_0^t \sigma(s,\omega,X_s(\omega))\,dW_s(\omega) \ . \tag{5}$$

We note that the previous definition gives for granted that the two integrals in (5) are well-defined. Now, we need a result that tells us when an SDE has some solutions.

Theorem 11. *Let us consider the SDE*

$$\begin{cases} dX_t(\omega) = b(t,\omega,X_t(\omega))\,dt + \sigma(t,\omega,X_t(\omega))\,dW_t(\omega) \ , \\ X_t(\omega) = \bar{x} \ . \end{cases}$$

*If the following conditions hold:*

- *for all $\bar{\omega} \in \Omega$ the functions $b(t,\bar{\omega},x), \sigma(t,\bar{\omega},x)$ are measurable;*
- *for all $\bar{x} \in \mathbb{R}$ the processes $b(t,\omega,\bar{x}), \sigma(t,\omega,\bar{x})$ are progressively measurable;*
- *there exists a constant $L > 0$ such that for all $\bar{t} \in [0,T]$ and for all $\bar{\omega} \in \Omega$ the following conditions hold $\forall x, y \in \mathbb{R}^n$*

$$\begin{aligned} \|b(\bar{t},\bar{\omega},x) - b(\bar{t},\bar{\omega},y)\| &\leq L\|x-y\| \ , \\ \|\sigma(\bar{t},\bar{\omega},x) - \sigma(\bar{t},\bar{\omega},y)\| &\leq L\|x-y\| \ ; \end{aligned} \tag{6}$$

- *the processes* $b(t,\omega,0)$ *and* $\sigma(t,\omega,0)$ *belong to the process space* $\mathcal{L}^2$*;*

*then there exists one and only one process* $X_t(\omega) \in \mathcal{C}$ *which is a solution of the stochastic differential equation. Moreover, the solution process satisfies the inequality*

$$\mathbb{E}\left(\max_{t\in[0,T]} \|X_t(\omega)\|^p\right) \leq K\left(1+\|x\|^p\right) \ ,$$

*where* $K>0$*, and* $p=1,2$*.*

*Proof.* [**28**, p.49], [**9**, Appendix D, p.397] □

A particular case is constituted by linear one-dimensional SDEs. Now, we study the following SDE (which is called linear SDE)

$$\left\{\begin{array}{l} dX_t(\omega) = b_t X_t(\omega)\,dt + \sigma_t X_t(\omega)\,dW_t(\omega) \ , \\ X_t(\omega) = \bar{x} \ , \end{array}\right. \tag{7}$$

where $b_t, \sigma_t \in C^0([0,T],\mathbb{R})$ are continuous deterministic functions. It is easy to prove that the functions $b(t,\omega,x) \triangleq b_t x$ and $\sigma(t,\omega,x) \triangleq \sigma_t x$ satisfy the hypothesis (6), therefore there exists one and only one process $X_t(\omega) \in \mathcal{C}$ which is a solution of the SDE (7). In order to find such a process $X_t(\omega)$, we consider (in analogy with the deterministic case solution) the process $\exp(Y_t(\omega))$ (where $Y_t(\omega)$ is an Itô process). By the chain rule,

$$\begin{aligned} & d\left[\exp\left(Y_t(\omega)\right)\right] = \\ & = \exp\left(Y_t(\omega)\right) dY_t(\omega) + \tfrac{1}{2}\exp\left(Y_t(\omega)\right)\left(Y_t^{\text{diff}}(\omega)\right)^2 dt \ , \\ & = \exp\left(Y_t(\omega)\right)\left\{\left[Y_t^{\text{drift}}(\omega) + \tfrac{1}{2}\left(Y_t^{\text{diff}}(\omega)\right)^2\right] dt + Y_t^{\text{diff}}(\omega)\,dW_t(\omega)\right\} \ . \end{aligned}$$

The process $\bar{x}\exp(Y_t(\omega))$ is a solution of (7) if and only if

$$\left\{\begin{array}{l} Y_t^{\text{diff}}(\omega) = \sigma_t \ , \\ Y_t^{\text{drift}}(\omega) = b_t - \sigma_t^2/2 \ . \end{array}\right.$$

Therefore, the process $(t,\omega) \mapsto \exp\left(\int_0^t (b_s - \sigma_s^2/2)\,ds + \int_0^t \sigma_s dW_s(\omega)\right)$ plays an important role in the solution of linear SDEs. This process is called *stochastic exponential* and has some important properties. First of all, let us define

$$\eta_t[b,\sigma](\omega) \triangleq \exp\left(\int_0^t \left(b_s(\omega) - \sigma_s^2(\omega)/2\right) ds + \int_0^t \sigma_t(\omega)\,dW_s(\omega)\right) \ ,$$

where $b_s(\omega) \in \mathcal{L}^1$, and $\sigma_t(\omega) \in \mathcal{L}^2$. Clearly, a continuous deterministic process as $b_t$ or $\sigma_t$ belongs to the space $\mathcal{L}^2$, therefore $\bar{x}\ \eta_t[b,\sigma](\omega)$ is well-defined and is the solution of the SDE (7). Moreover, we can prove

[**14**, p.73] that every positive one-dimensional Itô process $S_t(\omega)$ can be written, using the stochastic exponential, as

$$S_t(\omega) = S_0(\omega)\ \eta_t[b_t, \sigma_t]$$

for some $b_t(\omega) \in \mathcal{L}^1$, and $\sigma_t(\omega) \in \mathcal{L}^2$. Therefore, the stochastic exponential is the first process to study in order to model positive quantities. Some useful properties of the stochastic exponential are:

$$\begin{aligned} &1/\eta_t[b, \sigma](\omega) = \eta_t[-b + \sigma^2, -\sigma](\omega) \ , \\ &\eta_t[b, \sigma](\omega)\ \eta_t[B, \Sigma](\omega) = \eta_t[b + B + \sigma\Sigma, \sigma + \Sigma](\omega) \ . \end{aligned}$$

These relations may be directly obtained by using the chain rule of the stochastic calculus.

## 2. Backward Stochastic Differential Equations

In this Section, we introduce the concept of backward stochastic differential equation (BSDE for short) and we define the form that a solution of such an equation must have. A non-linear BSDE on the standard filtered probability space $\left(\Omega, \mathcal{E}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ is defined as follows:

$$\left\{ \begin{array}{l} dY_t(\omega) = h(t, \omega, Y_t(\omega), Z_t(\omega))\, dt + Z_t(\omega)\, dW_t(\omega) \ , \\ Y_T(\omega) = \vartheta(\omega) \ , \end{array} \right. \tag{8}$$

where $h(t, \omega, y, z) : [0, T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^{n\times m} \to \mathbb{R}^n$, $\vartheta(\omega)$ is a fixed $n$-dimensional random variable, and $W_t(\omega)$ is an $m$-dimensional Wiener process.

Definition 7. *Two processes $Y_t(\omega) \in \mathcal{C}$ and $Z_t(\omega) \in \mathcal{L}^2$ are a solution of the BSDE (8) if and only if*

$$Y_t(\omega) = \vartheta(\omega) - \int_t^T h(t, \omega, Y_s(\omega), Z_s(\omega))\, ds - \int_t^T Z_s(\omega)\, dW_s(\omega) \ . \tag{9}$$

We note that the above definition requires that the two integrals in (9) are well-defined. As with the SDEs, we need a result that tells us when a BSDE has some solutions.

Theorem 12. *Let us consider the following BSDE*

$$\left\{ \begin{array}{l} dY_t(\omega) = h(t, \omega, Y_t(\omega), Z_t(\omega))\, dt + Z_t(\omega)\, dW_t(\omega) \ , \\ Y_T(\omega) = \vartheta(\omega) \ . \end{array} \right.$$

*If the following conditions hold*

- $\vartheta(\omega) \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$;

- *for all* $\bar{\omega} \in \Omega$ *the function* $h(t, \bar{\omega}, x, y) : [0,T] \times \mathbb{R}^n \times \mathbb{R}^{n \times m} \to \mathbb{R}^n$ *is measurable;*
- *for all* $\bar{x}, \bar{y} \in \mathbb{R}$ *the process* $h(t, \omega, \bar{x}, \bar{y}) : [0,T] \times \Omega \to \mathbb{R}^n$ *is progressively measurable;*
- *there exists a constant* $L > 0$ *such that for all* $\bar{t} \in [0,T]$, *and for all* $\bar{\omega} \in \Omega$ *the following condition holds* $\forall x, y, v, w \in \mathbb{R}$*:*

$$\|h(\bar{t}, \bar{\omega}, x, y) - h(\bar{t}, \bar{\omega}, v, w)\| \leq L(\|x - v\| + \|y - w\|) \ ;$$

- *the process* $h(t, \omega, 0, 0) : [0,T] \times \Omega \to \mathbb{R}^n$ *belongs to the process space* $\mathcal{L}^2$*;*

*then there exists a unique pair of processes* $(Y_t(\omega), Z_t(\omega)) \in \mathcal{C} \times \mathcal{L}^2$ *that is a solution of the BSDE.*

*Proof.* [**28**, p.355] □

A one-dimensional linear backward stochastic differential equation (LBSDE for short) on the filtered probability space $\left(\Omega, \mathcal{E}, \{\mathcal{F}_t\}_{t \in [0,T]}, \mathbb{P}\right)$ is defined as follows:

$$\left\{ \begin{array}{l} dY_t(\omega) = (b_t(\omega) Y_t(\omega) + \sigma_t(\omega) Z_t(\omega)) \, dt + Z_t(\omega) \, dW_t(\omega) \ , \\ Y_T(\omega) = \vartheta(\omega) \ , \end{array} \right. \tag{10}$$

where $b_t(\omega), \sigma_t(\omega)$ are progressively measurable processes, $\vartheta(\omega) \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, and $W_t(\omega)$ is a one-dimensional Wiener process. We can prove [**28**, p.349] that if the processes $b_t(\omega), \sigma_t(\omega)$ belong to the processes space $\mathcal{L}^2$ and if they are essentially bounded, then there exists a unique pair of processes $(Y_t(\omega), Z_t(\omega)) \in \mathcal{C} \times \mathcal{L}^2$ that is a solution of the LBSDE (10). Moreover, the two processes $Y_t(\omega), Z_t(\omega)$ satisfy the following relation: there exists $K > 0$ such that

$$\mathbb{E}\left(\sup_{t \in [0,T]} |Y_t(\omega)|^2\right) + \mathbb{E}\left(\int_0^T |Z_t(\omega)|^2 \, dt\right) \leq K \ \mathbb{E}\left(|\vartheta(\omega)|^2\right) \ . \tag{11}$$

As previously done for the LSDE, we want to characterize the pair of processes $(Y_t(\omega), Z_t(\omega)) \in \mathcal{C} \times \mathcal{L}^2$ that solve the LBSDE (10). If $Y_t(\omega)$ is a process that satisfies (10), then[1]

$$\begin{aligned} & d\{\eta_t[B, \Sigma] Y_t\} = \\ & = Y_t d\eta_t[B, \Sigma] + \eta_t[B, \Sigma] dY_t + \eta_t[B, \Sigma] \Sigma_t Z_t dt \ , \\ & = \eta_t[B, \Sigma] B_t Y_t dt + \eta_t[B, \Sigma] \Sigma_t Y_t dW_t + \\ & \qquad + \eta_t[B, \Sigma](b_t Y_t + \sigma_t Z_t) \, dt + \eta_t[B, \Sigma] Z_t dW_t \\ & \qquad + \eta_t[B, \Sigma] \Sigma_t Z_t dt \ . \end{aligned}$$

[1] In order to have a simpler notation, we do not indicate the dependence on $\omega$ in the stochastic processes.

The drift of the process $\eta_t\,[B,\Sigma]\;\;Y_t$ is

$$\eta_t\,[B,\Sigma]\;\;\{(B_t+b_t)\,Y_t+(\sigma_t+\Sigma_t)\,Z_t\}\;\;.$$

Therefore, if we choose $B_t=-b_t$ and $\Sigma_t=-\sigma_t$, then we obtain a process with zero-drift. Hence, for all $\bar{t}\in[0,T]$

$$\eta_T\,[-b,-\sigma]\;\;\vartheta-\eta_{\bar{t}}\,[-b,-\sigma]\;\;Y_{\bar{t}}=\int_{\bar{t}}^{T}(\cdots)\,dW_t\;\;.$$

If we apply the conditional expectation under the $\sigma$-field $\mathcal{F}_{\bar{t}}$ we get

$$\mathbb{E}\left(\eta_T\,[-b,-\sigma]\;\;\vartheta\right|\mathcal{F}_{\bar{t}})=\eta_{\bar{t}}\,[-b,-\sigma]\;\;Y_{\bar{t}}\;\;. \tag{12}$$

As $\bar{t}\in[0,T]$ was chosen arbitrarily, this equation holds for all $\bar{t}\in[0,T]$. The process

$$(t,\omega)\mapsto\mathbb{E}\left(\eta_T\,[-b,-\sigma]\;\;\vartheta\right|\mathcal{F}_{\bar{t}})$$

is a continuous square integrable martingale [**28**, p.351]. Therefore, by the martingale representation theorem, there exists a process $H_t\,(\omega)\in\mathcal{H}^2$ such that

$$\begin{aligned}&\mathbb{E}\left(\eta_T\,[-b,-\sigma]\;\;\vartheta\right|\mathcal{F}_t)=\\&=\mathbb{E}\left(\eta_T\,[-b,-\sigma]\;\;\vartheta\right)+\textstyle\int_0^t H_s dW_s\;\;,\\&=\bar{y}+\textstyle\int_0^t H_s dW_s\;\;.\end{aligned}$$

Substituting this relation in (12) we obtain,

$$Y_t=\eta_t\left[b+\sigma^2,\sigma\right]\;\;\left(\bar{y}+\int_0^t H_s dW_s\right)\;\;.$$

This relation gives us the form of the process $Y_t\,(\omega)$ (it is an implicit form because we do not know the process $H_t\,(\omega)$ explicitly). Therefore we define

$$\begin{aligned}&Y_t\triangleq\eta_t\,[b+\sigma^2,\sigma]\;\;\left(\bar{y}+\textstyle\int_0^t H_s dW_s\right)\;\;,\\&Z_t\triangleq Y_t\sigma_t+\eta_t\,[b+\sigma^2,\sigma]\;\;H_t\;\;.\end{aligned}$$

Now we want to prove that the processes $Y_t\,(\omega)$ and $Z_t\,(\omega)$ just defined are a solution of the equation (10). By differentiating the process $Y_t\,(\omega)$

we obtain

$$\begin{aligned}
d\left[Y_t\right] &= d\left[\eta_t\left[b+\sigma^2,\sigma\right]\ \left(\bar{y}+\int_0^t H_s dW_s\right)\right]\ , \\
&= \left(\bar{y}+\int_0^t H_s dW_s\right)\ d\eta_t\left[b+\sigma^2,\sigma\right]+\eta_t\left[b+\sigma^2,\sigma\right]\ H_t dW_t+ \\
&\qquad +\eta_t\left[b+\sigma^2,\sigma\right]\ \sigma_t H_t dt\ , \\
&= Y_t\left(b_t+\sigma_t^2\right)dt+Y_t\sigma_t dW_t+\eta_t\left[b+\sigma^2,\sigma\right]\ H_t dW_t+ \\
&\qquad +\eta_t\left[b+\sigma^2,\sigma\right]\ \sigma_t H_t dt\ , \\
&= Y_t b_t dt+\sigma_t\left(\sigma_t Y_t+\eta_t\left[b+\sigma^2,\sigma\right]\ H_t\right)dt+Z_t dW_t\ , \\
&= \left(Y_t b_t+\sigma_t Y_t\right)dt+Z_t dW_t\ .
\end{aligned}$$

We note that, if the relation (11) holds, then the uniqueness of the processes $Y_t(\omega)$ and $Z_t(\omega)$ follows directly from the linearity of the equation.

CHAPTER 3

# Stochastic Control Problems

## 1. Formulation

Given the filtered probability space $\left(\Omega, \mathcal{E}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$, we consider a controlled SDE:

$$
(13) \qquad \begin{cases} dX_t = b\,(t, X_t, u_t)\,dt + \sigma\,(t, X_t, u_t)\,dW_t \ , \\ X_0 = x \ , \end{cases}
$$

where $b\,(t,x,u) : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$, $\sigma\,(t,x,u) : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^{n\times m}$, and $W_t\,(\omega) = (W_t^1\,(\omega), ..., W_t^m\,(\omega))'$ is an $m$-dimensional Wiener process. The function $u_t\,(\omega)$ is called the control and it represents the action/decision/policy of the decision-maker. At all times $t \in [0,T]$, the controller is knowledgeable about some information (as specified in the information structure $\{\mathcal{F}_t\}_{t\in[0,T]}$) about what has happened up to the time $t$, but she/he does not know what is going to happen afterwards, due to the uncertainty of the system. As a consequence, for all times $t \in [0,T]$, the controller does not exercise her/his decision $u_t\,(\omega)$ before the time $t$ really comes; in this way she/he can use all the available information in order to obtain the best result. This non-anticipative restriction can be represented, in mathematical terms, by the following constraint: "the process $u_t\,(\omega)$ is adapted to the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$" [**28**, p.63].

The set of all possible actions is $U$ (for our purposes it is a convex subset of $\mathbb{R}^k$), hence the control is a process $u_t\,(\omega) : [0,T]\times\Omega \to U$. The decision-maker chooses the action $u_t\,(\omega)$ in order to maximize her/his utility. Therefore, we introduce a cost functional that describes the preference of the decision-maker:

$$
(14) \qquad J\,[u_t\,(\omega)] \triangleq \mathbb{E}\left(\int_0^T f\,(t, X_t\,(\omega), u_t\,(\omega))\,dt + h\,(X_T\,(\omega))\right) \ .
$$

Generally, the functions $f\,(x,u) : [0,T]\times\mathbb{R}^n\times U \to \mathbb{R}$ and $h\,(x) : \mathbb{R}^n \to \mathbb{R}$ are chosen in such a way that the integral in (14) exists (they may be not finite, but they must be well-defined). Sometimes, we require that these functions are continuous (or differentiable) and have a polynomial growth in the variable $x$: i.e. for all $t \in [0,T]$, $x \in \mathbb{R}^n$, and $u \in U$ the relations

$$\|f(t,x,u)\| \le L(1+\|x\|+\|u\|)^p \ ,$$
$$\|h(x)\| \le L(1+\|x\|)^p$$

hold for some positive real numbers $p, L$.

The set where the control processes are chosen is called the admissible control set and it is denoted by $\mathcal{U}$. It is defined as follows:

**control-state link:** a class of almost everywhere identical measurable and adapted processes $u_t^*(\omega)$ is in the set $\mathcal{U}$ if and only if there exists a unique process $X_t^*(\omega) \in \mathcal{C}$ which is a solution of the SDE (13) where the chosen control is $u_t^*(\omega)$.

A set of hypotheses that assure the good-definition of the control-state link is explicitly written in the following definition.

DEFINITION 8. *A stochastic control problem is called linear if*

- $\mathcal{U} = \mathcal{L}^2$;
- $b(t,x,u) = A_t x + B_t u$, $\sigma(t,x,u) = (\sigma^1(t,x,u), ..., \sigma^m(t,x,u))$ *and* $\sigma^i(t,x,u) = C_t^i x + D_t^i u$ *for all* $i = 1, ..., m$ *where* $A_t$, $B_t$, $C_t^i$, $D_t^i$ *are continuous and deterministic functions from* $[0,T]$ *to the following space:* $A_t, C_t^i \in \mathbb{R}^{n\times n}$, $B_t, D_t^i \in \mathbb{R}^{k\times n}$, *then the (13) becomes:*

$$\begin{cases} dX_t = (A_t X_t + B_t u_t)\,dt + \sum_{i=1}^m (C_t^i X_t + D_t^i u_t)\,dW_t^i \ , \\ X_0 = x \ ; \end{cases}$$

- *the functions* $f(t,x,u) : [0,T]\times\mathbb{R}^n\times U \to \mathbb{R}$ *and* $h(x) : \mathbb{R}^n \to \mathbb{R}$ *are twice differentiable with continuity in all their variables and have a polynomial growth in the variable* $x$.

A stochastic control problem can be formulated as follows: the goal is to find $u_t^*(\omega) \in \mathcal{U}$ (if it ever exists) such that

$$J[u_t^*(\omega)] = \inf_{u_t(\omega)\in\mathcal{U}} J[u_t(\omega)] \ .$$

We say that a stochastic control problem is *finite* if and only if

$$\inf_{u_t(\omega)\in\mathcal{U}} J[u_t(\omega)] > -\infty \ ;$$

we say that it is *solvable* if and only if there exists at least one control $u_t^*(\omega) \in \mathcal{U}$ such that

$$J[u_t^*(\omega)] = \inf_{u_t(\omega)\in\mathcal{U}} J[u_t(\omega)] \ .$$

If a stochastic control problem is *finite and solvable,* then there exists at least one control $u_t^*(\omega) \in \mathcal{U}$ such that

$$J[u_t^*(\omega)] = \inf_{u_t(\omega)\in\mathcal{U}} J[u_t(\omega)] > -\infty \ ,$$

then $u_t^*(\omega)$ is called an *optimal control* and the corresponding state process $X_t^*(\omega)$ and the couple $(u_t^*(\omega), X_t^*(\omega))$ are called *optimal state* and *optimal pair*, respectively.

The first natural question is: "What kind of conditions have the functions $b, \sigma, f, h$, to satisfy in order to assure that the stochastic control problem is finite and solvable?" The following theorem gives two sets of conditions that answer to this question.

THEOREM 13. *Let us consider a linear stochastic control problem with the motion equation*

$$\begin{cases} dX_t = (AX_t + Bu_t)\,dt + (CX_t + Du_t)\,dW_t \ , \\ X_0 = x \ , \end{cases}$$

*where $A, B, C, D$ are constant matrices of suitable size, and $W_t(\omega)$ is a one-dimensional Wiener process. If one of the following hypotheses is satisfied, then there exists a control $u_t^*(\omega) \in \mathcal{U} = \mathcal{L}^2$ such that*

$$J[u_t^*(\omega)] = \inf_{u_t(\omega)\in\mathcal{U}} J[u_t(\omega)] > -\infty \ .$$

**H1:** Lower Bound Hypothesis

- *$U$ is convex and closed;*
- *the function $f$ does not depend explicitly on time, $f(t,x,u) = f(x,u)$;*
- *$f(x,u), h(x)$ are both convex functions;*
- *there exist two constants $L, M > 0$ such that for all $x \in \mathbb{R}^n$, $u \in U$*

$$\begin{aligned} f(x,u) &\geq M\|u\|^2 - L \ , \\ h(x) &\geq -L \ . \end{aligned}$$

**H2:** Compactness Hypothesis

- *$U$ is convex and compact;*
- *the function $f$ does not depend explicitly on time, $f(t,x,u) = f(x,u)$;*
- *$f(x,u), h(x)$ are both convex functions.*

*Proof.* [**28**, p.68]. □

## 2. Peng's Stochastic Maximum Principle

In this Section we introduce the necessary conditions for the optimality in a stochastic control problem. The following result is related to the linear stochastic control problem; it holds for a more general class of problems, but this formulation of the necessary conditions is useful for the problems we are dealing with.

THEOREM 14 (Peng's Maximum Principle). *Given a linear stochastic control problem, if there exists an optimal pair* $(u_t^*(\omega), X_t^*(\omega))$ *then there exist the processes* $p_t(\omega), P_t(\omega) \in \mathcal{C}$, *and* $q_t^i(\omega), Q_t^i(\omega) \in \mathcal{L}^2$ *with* $p_t(\omega), q_t^i(\omega) \in \mathbb{R}^n$, *and* $P_t(\omega), Q_t^i(\omega) \in \mathbb{S}^n$ *(i.e. the space of the symmetric matrices of dimension* $n \times n$*) for all* $i = 1, ..., m$, *such that*[1]

$$
(15) \qquad \begin{cases} dp_t = -\left[A_t' p_t + \sum_{i=1}^m (C_t^i)' q_t^i - \nabla_x f(t, X_t^*, u_t^*)\right] dt \\ \qquad\qquad + \sum_{i=1}^m q_t^i dW_t^i \ , \\ p_T = -\nabla_x h(X_T^*) \ , \end{cases}
$$

$$
\begin{cases} dP_t(\omega) = -[A_t' P_t + P_t A_t + \sum_{i=1}^m \left((C_t^i)' P_t C_t^i\right) + \\ \qquad\qquad + \sum_{i=1}^m (C_t^i)' Q_t^i + Q_t^i C_t^i - \\ \qquad\qquad \mathrm{Hess}_x \left(f(t, X_t^*, u_t^*)\right)] dt + \sum_{i=1}^m Q_t^i dW_t^i \ , \\ P_T = -\mathrm{Hess}_x \left(h(X_T^*)\right) \ . \end{cases}
$$

*Moreover, if we define the* $\mathcal{H}$*-function as*

$$
\begin{aligned}
&\mathcal{H}(t, x, u) = \\
&\langle (A_t x + B_t u), p_t \rangle - f(t, x, u) + \\
&+\tfrac{1}{2}\mathrm{tr}\left((C_t^1 x + D_t^1 u, ..., C_t^m x + D_t^m u)' P_t (C_t^1 x + D_t^1 u, ..., C_t^m x + D_t^m u)\right) \\
&+\mathrm{tr}\{ (C_t^1 x + D_t^1 u, ..., C_t^m x + D_t^m u)' [(q_t^1, ..., q_t^m)' \\
&-P_t (C_t^1 X_t^* + D_t^1 u_t^*, ..., C_t^m X_t^* + D_t^m u_t^*)]\} \ ,
\end{aligned}
$$

*then the following maximum condition holds (*$\lambda \otimes \mathbb{P}$*-almost everywhere):*

$$
\mathcal{H}(t, X_t^*, u_t^*) = \max_{w \in U} \mathcal{H}(t, X_t^*, w) \ .
$$

*Proof.* [**28**, p.118]. $\square$

The necessary conditions play the role of the equation $f'(x) = 0$ in the finite dimensional optimization problems. Hence, a set of processes that satisfies the necessary conditions, may not be a minimum point for the objective functional. Therefore, we need further conditions in order to assure that a set of processes that satisfies the necessary conditions singles out a minimum point of the objective functional. This is the role played by the sufficient conditions that follow.

THEOREM 15. *Let us consider a linear stochastic optimal control problem such that the functions* $f(t, x, u), h(x)$ *are twice continuously differentiable with respect to the variable* $x$, *and have a polynomial growth in their variables. Suppose that:*

- $h(x)$ *is convex;*

[1]In order to have a simpler notation, we do not indicate the dependence on $\omega$ in the stochastic processes.

- *the function*

$$\begin{aligned}(x,u) \mapsto & \left(A_t x + B_t u\right) p_t\left(\omega\right) - f\left(t,x,u\right) + \\ & + \operatorname{tr}\left(\left(q_t^1, ..., q_t^m\right)' \left(C_t^1 x + D_t^1 u, ..., C_t^m x + D_t^m u\right)\right)\end{aligned}$$

  *is concave for all* $t \in [0,T]$, $\mathbb{P}$*-almost everywhere.*

*Then, if there exists a set of processes that satisfies the necessary conditions, then the control-state pair connected with this set of processes is an optimal pair.*

*Proof.* [**28**, p.139]. $\square$

CHAPTER 4

# The LQ Control Problem

## 1. Introduction

In this Chapter we are dealing with a special case of stochastic control problem where, on one hand, the state equations are linear in both state and control, and on the other hand, the cost functions are quadratic (this kind of problems is called linear quadratic, LQ for short). The LQ problems are very important because many non-linear control problems can be reasonably approximated by the LQ problems. Moreover, the solution of the LQ problems has a simple structure, and can be characterized by the solution of a non-linear ordinary differential equation, called the Riccati equation.

First of all, let us introduce the structure of a LQ stochastic control problem. Given a linear controlled SDE

$$\begin{cases} dX_t = (A_t X_t + B_t u_t)\,dt + \sum_{i=1}^m (C_t^i X_t + D_t^i u_t)\,dW_t^i \ , \\ X_0 = x \end{cases}$$

we can choose the following operator as cost functional to minimize:

$$J\left[u_t(\omega)\right] = \mathbb{E}\left(\frac{1}{2}\int_0^T \left(\langle F_t X_t, X_t\rangle + \langle R_t u_t, u_t\rangle\right) dt + \frac{1}{2}\langle G X_T, X_T\rangle\right) \ ,$$

where $F_t, R_t$ are continuous matrix values functions such that $F_t \in C^0([0,T], \mathbb{S}^n)$, $R_t \in C^0\left([0,T], \mathbb{S}^k\right)$, while $G$ is a fixed matrix such that $G \in \mathbb{S}^n$. Now, let us apply the Maximum Principle to the LQ stochastic control problem written above.

Theorem 16 (Peng's Maximum Principle for LQ problems). *Given an LQ problem, if there exists an optimal pair $\left(u_t^*(\omega), X_t^*(\omega)\right)$, then there are $p_t(\omega), P_t(\omega) \in \mathcal{C}$, and $q_t^i(\omega), Q_t^i(\omega) \in \mathcal{L}^2$ with $p_t(\omega), q_t^i(\omega) \in \mathbb{R}^n$, and $P_t(\omega), Q_t^i(\omega) \in \mathbb{S}^n$ (i.e. the space of the symmetric matrices* *of dimension $n \times n$) for all $i = 1, ..., m$, such that*[1]

$$\begin{cases} dp_t = -\left[A_t' p_t + \sum_{i=1}^m \left(C_t^i\right)' q_t^i - F_t X_t^*\right] dt + \sum_{i=1}^m q_t^i dW_t^i \ , \\ p_T = -G X_T^* \ , \end{cases}$$

$$\begin{cases} dP_t(\omega) = -[A_t' P_t + P_t A_t + \sum_{i=1}^m \left(\left(C_t^i\right)' P_t C_t^i\right) + \\ \qquad\qquad + \sum_{i=1}^m \left(\left(C_t^i\right)' Q_t^i + Q_t^i C_t^i\right) - F_t] dt + \sum_{i=1}^m Q_t^i dW_t^i \ , \\ P_T = -G \ , \end{cases}$$

*and moreover $\lambda \otimes \mathbb{P}$-almost everywhere*

$$\begin{array}{l} R_t u_t^* - B_t' p_t - \sum_{i=1}^m \left(D_t^i\right)' q_t^i = 0 \ , \\ R_t + \sum_{i=1}^m \left(D_t^i\right)' P_t D_t^i \succeq 0 \quad \textit{i.e. positive semidefinite} \ . \end{array}$$

*Proof.* [**28**, p.309]. □

These conditions require to solve a forward-backward SDE, which is not a simple task. On the other hand, the conditions suggest [**28**, p.313] that $p_t = \pi_t X_t^*$ for some process $\pi_t$. Starting from this suggestion we arrive to the Riccati equation which permits us to characterize $\pi_t$ in order to solve the necessary conditions.

## 2. Riccati Equation

**2.1. Notation.** First of all, let us define more in detail the process space we are working on (in the following $p \in [1, +\infty)$). Let $L_{\mathcal{F}}^p\left(0, T; \mathbb{R}^k\right)$ be the quotient space, with respect to the equivalence relation of almost everywhere identity, of measurable processes $u_t(\omega)$ defined on the time interval $[0, T]$ which are $\{\mathcal{F}_t\}_{t \in [0,T]}$-adapted, $\mathbb{R}^k$-valued, and satisfy the following condition

$$\mathbb{E}\left(\int_0^T \|u_t(\omega)\|_{\mathbb{R}^k}^p\right) < +\infty \ .$$

It is a Banach space with the norm $\|u_t(\omega)\| = \mathbb{E}\left(\int_0^T \|u_t(\omega)\|_{\mathbb{R}^k}^p dt\right)^{1/p}$. Moreover, let $L_{\mathcal{F}}^p(\Omega; C([0, T]; \mathbb{R}^n))$ be the quotient space, with respect to the stochastic equivalence relation, of measurable and continuous processes $X_t(\omega)$ defined on the time interval $[0, T]$ which are $\{\mathcal{F}_t\}_{t \in [0,T]}$-adapted, $\mathbb{R}^n$-valued, and such that

$$\mathbb{E}\left(\sup_{t \in [0,T]} \|X_t(\omega)\|_{\mathbb{R}^n}^p\right) < +\infty \ .$$

[1] In order to have a simpler notation, we do not indicate the dependence on $\omega$ in the stochastic processes.

It is a Banach space with the norm

$$\|X_t(\omega)\| = \mathbb{E}\left(\sup_{t\in[0,T]} \|X_t(\omega)\|^p_{\mathbb{R}^n}\right)^{1/p} .$$

In the following we shall employ the usual convention of suppressing the $\omega$-dependence notation of all the stochastic processes.

The data of the problem define the motion equation and the objective function; they are:

- $A_t, C^i_t \in C^0([0,T];\mathbb{R}^{n\times n})$;
- $B_t, D^i_t \in C^0\left([0,T];\mathbb{R}^{n\times k}\right)$ for all $i = 1, ..., m$;
- $F_t \in C^0([0,T];\mathbb{S}^n)$ ($\mathbb{S}^n$ is the space of the symmetric $n \times n$ matrices);
- $R_t \in C^0\left([0,T];\mathbb{S}^k\right)$;
- $G \in \mathbb{S}^n$.

One can prove [**28**, p.49] that, for all fixed $u_t \in L^2_{\mathcal{F}}\left(0,T;\mathbb{R}^k\right)$, there exists a unique $X_t \in L^2_{\mathcal{F}}(\Omega; C([0,T];\mathbb{R}^n))$ such that

$$\begin{cases} dX_t = (A_t X_t + B_t u_t)\, dt + \sum_{i=1}^m (C^i_t X_t + D^i_t u_t)\, dW^i_t \ , \\ X_0 = x \ . \end{cases}$$

The LQ problem consists in finding a control $u_t \in L^2_{\mathcal{F}}\left(0,T;\mathbb{R}^k\right)$ that minimizes the functional

$$J[u_t] = \mathbb{E}\left(\frac{1}{2}\int_0^T (\langle F_t X_t, X_t\rangle + \langle R_t u_t, u_t\rangle)\, dt + \frac{1}{2}\langle G X_T, X_T\rangle\right) .$$

Under these assumptions we are not sure that the problem is well-defined, but we can provide necessary conditions in order to obtain this result. We note that the matrices $F_t, R_t, G$ are possibly indefinite (but generally in the literature it is required that all these matrices are positive definite in order to have a well-defined problem).

### 2.2. Riccati equation.

THEOREM 17 (Riccati equation). *Let us assume that there exists a matrix function* $\pi_t \in C^1([0,T];\mathbb{R}^{n\times n})$ *such that*

$$
(16)\qquad \begin{cases} \dot{\pi}_t + \pi_t A_t + A_t'\pi_t + \sum_{i=1}^{m} \left(C_t^i\right)' \pi_t C_t^i + F_t - \Phi_t'\Psi_t^{-1}\Phi_t = 0 \ , \\ \Phi_t \triangleq B_t'\pi_t + \sum_{i=1}^{m} \left(D_t^i\right)' \pi_t C_t^i \ , \\ \Psi_t \triangleq R_t + \sum_{i=1}^{m} \left(D_t^i\right)' \pi_t D_t^i \ , \\ \pi_T = G \ , \\ R_t + \sum_{i=1}^{m} \left(D_t^i\right)' \pi_t D_t^i \succ 0 \quad \textit{i.e. positive definite} \ , \end{cases}
$$

*therefore the optimal control can be written in a feedback form as*

$$
u_t^* = -\Psi_t^{-1}\Phi_t X_t^* \ .
$$

First of all, let us present some useful results that give some further information on the solution of the ODE introduced by the above theorem.

THEOREM 18. *Let* $\pi_t^*, \pi_t^o \in C^1([0,T];\mathbb{R}^{n\times n})$ *satisfy the conditions (16), then* $\pi_t^* = \pi_t^o$ *for all* $t \in [0,T]$.

*Proof.* Let us define $\varphi_t = \pi_t^* - \pi_t^o$; it can be proved [**28**, p.320] that

$$
\|\dot{\varphi}_t\| \leq K \ \|\varphi_t\| \ ,
$$

where $K > 0$. Hence, as $\varphi_T = 0$, we can use the Gronwall Lemma and obtain that $\varphi_t = 0$ for all $t \in [0,T]$. □

LEMMA 19. *Let* $\pi_t \in C^1([0,T];\mathbb{R}^{n\times n})$ *satisfy the conditions (16), then* $\pi_t \in \mathbb{S}^n$ *for all* $t \in [0,T]$.

*Proof.* We note that if $\pi_t \in C^1([0,T];\mathbb{R}^{n\times n})$ satisfies (16), then the matrix function $\pi_t'$ satisfies the same conditions, too. But, the conditions (16) characterize a unique matrix function, therefore $\pi_t' = \pi_t$ for all $t \in [0,T]$. □

Now, let us prove the result previously introduced.

*Proof of Theorem 17.* Let $(u_t, X_t)$ be a control-state pair for the LQ problem. We consider the process $\langle \pi_t X_t, X_t \rangle$; its stochastic differential is

$$\begin{aligned}
&d\left[\langle \pi_t X_t, X_t\rangle\right] = \\
&= \langle d\left[\pi_t X_t\right], X_t\rangle + \langle \pi_t X_t, dX_t\rangle + \\
&\qquad + \textstyle\sum_{i=1}^m \langle \pi_t\left(C_t^i X_t + D_t^i u_t\right), C_t^i X_t + D_t^i u_t\rangle\, dt\ , \\
&= \langle \dot{\pi}_t X_t, X_t\rangle\, dt + \langle \pi_t dX_t, X_t\rangle + \langle \pi_t X_t, dX_t\rangle \\
&\qquad + \textstyle\sum_{i=1}^m \langle \pi_t\left(C_t^i X_t + D_t^i u_t\right), C_t^i X_t + D_t^i u_t\rangle\, dt\ .
\end{aligned}$$

We can rewrite the second and third term as

$$\begin{aligned}
\langle \pi_t dX_t, X_t\rangle &= \langle \pi_t\left(A_t X_t + B_t u_t\right), X_t\rangle\, dt \\
&\qquad + \textstyle\sum_{i=1}^m \langle \pi_t\left(C_t^i X_t + D_t^i u_t\right), X_t\rangle\, dW_t^i\ , \\
\langle \pi_t X_t, dX_t\rangle &= \langle \pi_t X_t, \left(A_t X_t + B_t u_t\right)\rangle\, dt \\
&\qquad + \textstyle\sum_{i=1}^m \langle \pi_t X_t, \left(C_t^i X_t + D_t^i u_t\right)\rangle\, dW_t^i\ .
\end{aligned}$$

Moreover, the argument of the sum in the fourth term has the following form:

$$\begin{aligned}
&\langle \pi_t\left(C_t^i X_t + D_t^i u_t\right), C_t^i X_t + D_t^i u_t\rangle = \\
&= \langle \pi_t C_t^i X_t + \pi_t D_t^i u_t, C_t^i X_t + D_t^i u_t\rangle\ , \\
&= \langle \pi_t C_t^i X_t, C_t^i X_t + D_t^i u_t\rangle + \langle \pi_t D_t^i u_t, C_t^i X_t + D_t^i u_t\rangle\ , \\
&= \left\langle \left(C_t^i\right)' \pi_t C_t^i X_t, X_t\right\rangle + 2\left\langle \left(D_t^i\right)' \pi C_t^i X_t, u_t\right\rangle + \left\langle \left(D_t^i\right)' \pi_t D_t^i u_t, u_t\right\rangle\ .
\end{aligned}$$

Hence, setting $\Pi_t$ as the drift of the process $\langle \pi_t X_t, X_t\rangle$, i.e.

$$\mathbb{E}\left(\langle \pi_T X_T, X_T\rangle\right) = \langle \pi_0 x, x\rangle + \mathbb{E}\left(\int_0^T \Pi_t dt\right)\ ;$$

we have that

$$\begin{aligned}
&\mathbb{E}\left(\langle G X_T, X_T\rangle\right) + \mathbb{E}\left(\textstyle\int_0^T \left(\langle F_t X_t, X_t\rangle + \langle R_t u_t, u_t\rangle\right) dt\right) = \\
&\langle \pi_0 x, x\rangle + \mathbb{E}\left(\textstyle\int_0^T \Pi_t dt\right) + \mathbb{E}\left(\textstyle\int_0^T \left(\langle F_t X_t, X_t\rangle + \langle R_t u_t, u_t\rangle\right) dt\right)\ ,
\end{aligned}$$

therefore

$$J\left[u_t\right] = \frac{1}{2}\langle \pi_0 x, x\rangle + \frac{1}{2}\mathbb{E}\left(\int_0^T \left(\Pi_t + \langle F_t X_t, X_t\rangle + \langle R_t u_t, u_t\rangle\right) dt\right)\ .$$

Now, we want to investigate the form of the process $\Pi_t + \langle F_t X_t, X_t\rangle + \langle R_t u_t, u_t\rangle$. By expanding the terms we obtain

$$\begin{aligned}
&\Pi_t + \langle F_t X_t, X_t\rangle + \langle R_t u_t, u_t\rangle = \\
&= \langle \dot{\pi}_t X_t, X_t\rangle + \langle \pi_t A_t X_t, X_t\rangle + \langle B_t' \pi_t X_t, u_t\rangle + \\
&\qquad + \langle A_t' \pi_t X_t, X_t\rangle + \langle B_t' \pi_t X_t, u_t\rangle + \left\langle \textstyle\sum_{i=1}^m \left(C_t^i\right)' \pi_t C_t^i X_t, X_t\right\rangle \\
&\qquad + 2\left\langle \textstyle\sum_{i=1}^m \left(D_t^i\right)' \pi C_t^i X_t, u_t\right\rangle + \left\langle \textstyle\sum_{i=1}^m \left(D_t^i\right)' \pi_t D_t^i u_t, u_t\right\rangle \\
&\qquad + \langle F_t X_t, X_t\rangle + \langle R_t u_t, u_t\rangle\ .
\end{aligned}$$

After recalling that

$$\dot{\pi}_t = -\pi_t A_t - A_t'\pi_t - \sum_{i=1}^{m} \left(C_t^i\right)' \pi_t C_t^i - F_t + \Phi_t' \Psi_t^{-1} \Phi_t \ ,$$

we get

$$\begin{aligned} &\Pi_t + \langle F_t X_t, X_t \rangle + \langle R_t u_t, u_t \rangle = \\ &= \left\langle \Phi_t' \Psi_t^{-1} \Phi_t X_t, X_t \right\rangle + 2 \langle \Phi_t X_t, u_t \rangle + \langle \Psi_t u_t, u_t \rangle \ . \end{aligned}$$

We know that $\Psi_t$ is a symmetric positive definite matrix, therefore there exists a symmetric matrix $\Psi_t^{1/2}$ such that $\Psi_t^{1/2}\Psi_t^{1/2} = \Psi_t$. Hence

$$\begin{aligned} &\left\langle \Phi_t' \Psi_t^{-1} \Phi_t X_t, X_t \right\rangle + 2 \langle \Phi_t X_t, u_t \rangle + \langle \Psi_t u_t, u_t \rangle = \\ &= \left\langle \Psi_t^{-1/2} \Phi_t X_t, \Psi_t^{-1/2} \Phi_t X_t \right\rangle + 2 \langle \Phi_t X_t, u_t \rangle + \left\langle \Psi_t^{1/2} u_t, \Psi_t^{1/2} u_t \right\rangle \ , \\ &= \left\langle \Psi_t^{-1/2} \Phi_t X_t + \Psi_t^{1/2} u_t, \Psi_t^{-1/2} \Phi_t X_t + \Psi_t^{1/2} u_t \right\rangle \ , \\ &= \left\| \Psi_t^{-1/2} \Phi_t X_t + \Psi_t^{1/2} u_t \right\|^2 \ . \end{aligned}$$

By substituting this result in the previous equation connected with the objective functional, we obtain

$$J\left[u_t\right] = \frac{1}{2} \langle \pi_0 x, x \rangle + \frac{1}{2} \mathbb{E} \left( \int_0^T \left\| \Psi_t^{-1/2} \Phi_t X_t + \Psi_t^{1/2} u_t \right\|^2 dt \right) \ .$$

This relation holds for all $u_t \in L_{\mathcal{F}}^2 \left(0, T; \mathbb{R}^k\right)$, therefore the optimal control can be characterized by choosing

$$u_t^* = -\Psi_t^{-1} \Phi_t X_t^* \ .$$

□

The main difference between deterministic and stochastic LQ problems is strictly connected with the form of the term

$$\Psi_t = R_t + \sum\nolimits_{i=1}^{m} \left(D_t^i\right)' \pi_t D_t^i \ .$$

If the problem is deterministic (or if the control does not affect the diffusion term, i.e. $D_t^i = 0$ for all $i = 1, ..., m$), then the theorem above holds if and only if $R_t \succ 0$. On the other hand, when $D_t^i \neq 0$ for some $i \in \{1, ..., m\}$, the result is useful even if $R_t$ is indefinite. This interesting observation appeared for the first time in the paper [**7**].

There is another problem connected with the theory of the stochastic Riccati equation. First of all, we recall that a necessary condition for both the solvability and the finiteness of a deterministic LQ problem is the condition $R_t \succeq 0$ [**28**, p.290]. On the other hand, this condition is not sufficient for the solvability, nor for the finiteness. On the other hand, if we assume that $R_t \succ 0$, then a deterministic LQ problem is finite and solvable if and only if the deterministic Riccati equation

$$\begin{cases} \dot{\pi}_t + \pi_t' A_t + A_t' \pi_t + F_t - \left(B_t' \pi_t\right)' R_t^{-1} B_t' \pi_t = 0 \ , \\ \pi_T = G \ , \\ R_t \succ 0 \end{cases}$$

has a solution[2] defined in the whole interval $[0,T]$ [**28**, p.296]. This quick summary on deterministic LQ problems, leaves us with the suggestion that if in the stochastic LQ problem one required that $R_t + \sum_{i=1}^m \left(D_t^i\right)' \pi_t D_t^i$ should be positive definite, then the existence of a solution of the stochastic Riccati equation would be equivalent to the solvability of a stochastic LQ problem. However, this is not true. In fact, it has been proved in the paper [**7**] that there exist some solvable stochastic LQ problems that do not satisfy the condition that $\Psi_t$ is positive definite. Recently, in the work [**1**], a generalized Riccati equation has been proposed, whose solution is equivalent to the solvability of a stochastic LQ problem. In the final part of the Section we present the more interesting results connected with this theory.

**2.3. Generalized Riccati equation.** We know that given a matrix $M \in \mathbb{R}^{m\times n}$ there exists a unique matrix $M^\dagger \in \mathbb{R}^{n\times m}$, called the pseudo inverse of the matrix $M$, such that

$$\begin{array}{ll} MM^\dagger M = M \ , & M^\dagger M M^\dagger = M^\dagger \ , \\ \left(MM^\dagger\right)' = MM^\dagger \ , & \left(M^\dagger M\right)' = M^\dagger M \ . \end{array}$$

We use the pseudo inverse in the following result.

Theorem 20 (Generalized Riccati equation). *Let us assume that there exists a matrix function $\pi_t \in C^1\left([0,T];\mathbb{R}^{n\times n}\right)$ such that*

$$\begin{cases} \dot{\pi}_t + \pi_t' A_t + A_t' \pi_t + \sum_{i=1}^m \left(C_t^i\right)' \pi_t C_t^i + F_t - \Phi_t' \Psi_t^\dagger \Phi_t = 0 \ , \\ \pi_T = G \ , \\ \Psi_t \Psi_t^\dagger \Phi_t - \Phi_t = 0 \ , \\ \Psi_t \succeq 0 \ , \end{cases} \tag{17}$$

*therefore the set of all optimal controls can be written in feedback form as*

$$u_t^* = -\left(\Psi_t^\dagger \Phi_t + Y_t - \Psi_t^\dagger \Psi_t Y_t\right) X_t^* + Z_t - \Psi_t^\dagger \Psi_t Z_t \ ,$$

[2]Moreover, if the LQ problem is standard (i.e. if not only $R_t$ positive definite, but also $G$ positive semidefinite) then the deterministic Riccati equation has a solution defined in $[0,T]$, hence the deterministic LQ problem is solvable. However, there exist solvable deterministic LQ problems that are not standard, hence being standard is not a necessary condition for the solvability.

*where* $Y_t \in L^2_{\mathcal{F}}\left(0,T;\mathbb{R}^{k\times n}\right)$ *and* $Z_t \in L^2_{\mathcal{F}}\left(0,T;\mathbb{R}^{k}\right)$.

*Proof.* [**1**, Theorem 3.1]. □

We know that the pseudo inverse coincides with the inverse when the matrix is non-singular. Hence, two particular cases of the Theorem above are the following:

- if for all $t \in [0,T]$ the matrix $\Psi_t \succ 0$, then the Riccati equation is the same as the one written in (16);
- if for all $t \in [0,T]$ the matrix $\Psi_t = 0$, then the Riccati equation is
$$\begin{cases} \dot{\pi}_t + \pi_t A_t + A_t' \pi_t + \sum_{i=1}^{m} \left(C_t^i\right)' \pi_t C_t^i + F_t = 0 \ , \\ \pi_T = G \ , \\ \Phi_t = 0 \ , \end{cases}$$
and all the controls in $L^2_{\mathcal{F}}\left(0,T;\mathbb{R}^{k}\right)$ are optimal.

We close this Section with a very interesting result that justifies the previous theorem. In fact, the relation above is not only a sufficient condition for the optimality, it is also a necessary one.

Theorem 21. *A stochastic LQ problem has an optimal control if and only if the generalized Riccati equation (17) has a solution.*

*Proof.* [**1**, Theorem 5.3]. □

CHAPTER 5

# A Marketing Model

## 1. Nerlove and Arrow's Model

One of the earliest studies in advertising is a paper written by Nerlove and Arrow [**15**]. In that work, the authors consider advertising as an investment in a stock (the goodwill $A_t$), which summarizes the effects of current and past advertising flow $a_t$; they assume that the goodwill evolution satisfies the first order linear differential equation

$$\dot{A}_t = -\delta A_t + a_t \ . \tag{18}$$

Then the goodwill decreases spontaneously with decay coefficient $\delta > 0$. $A_t$ is always positive and is sustained by the advertising investment $a_t$. The original model is introduced considering the long run scenario. However, the finite horizon hypothesis is already studied and it is quite simple to introduce: the firm wants to choose a price $p_t > 0$ and an advertising flow $a_t \geq 0$ in order to maximize its profit:

$$\int_0^T \left\{ S\left(A_t, p_t, Z_t\right) \left[ p_t - C\left(S\left(A_t, p_t, Z_t\right)\right)\right] - c\left(a_t\right) \right\} e^{-rt} \, dt + u\left(A_T\right) \ ,$$

where

- $S\left(A_t, p_t, Z_t\right)$ is the sales function, it is increasing in the variable $A$ (the more the goodwill is, the more the sales are), while it is decreasing in the variable $p$ (the more the price is, the less the sales are), and it depends on some exogenous factors out of the firm control, $Z_t$;
- $C\left(S\left(A_t, p_t, Z_t\right)\right)$ represents the production cost connected to the sales level $S\left(A_t, p_t, Z_t\right)$;
- $c\left(a_t\right)$ is the advertising cost that the firm has to sustain in order to obtain an advertising flow $a_t$;
- $r > 0$ is the discount rate;
- $u\left(A_T\right)$ is the utility considered as a prospective profit given by a final level of goodwill $A_T$ and it is an increasing function.

As already studied in [**5**], this model can also be useful in the problem of a new product introduction. In fact, the motion equation remains the same, but the objective functional can be modified as follows:

$$\underset{a_t \geq 0}{\text{Maximize}} \quad \int_0^T -c\,(a_t)\, e^{-rt}\; dt + u\,(A_T) \;\; .$$

During the pre-launch interval $[0, T]$ the firm can only present its product to the consumers, hence, it has only to sustain the advertising cost and does not have any revenues. At the time $T$, in some cases (concert, workshop,...) the firm obtains the total revenue, in other cases it starts to obtain the revenue which depends on the goodwill at the final time. In this situation $u\,(A_T)$ represents an estimate of the expected revenue. Therefore, all the utility obtained by the firm is concentrated at the end of the programming interval and it is described by the function $u\,(A_T)$. If the pre-launch interval is not too wide it is convenient to choose $r = 0$ and consider only the trade-off between advertising cost and final utility. An interesting example which can be described with this model is the following: we can consider an organization which has planned a social event (concert, workshop, soccer match), at a fixed time $T$. An advertising campaign has to be organized in order to stimulate as wide a participation as possible. We suppose that the customers buy the tickets only at time $T$ and that the demand depends on the event goodwill at $T$.

## 2. Advertising and a Stochastic Goodwill Process

While introducing the concept of goodwill we have assumed that the advertising flow is an investment in the stock of goodwill. Nowadays, the price evolution of the stocks is described using the stochastic processes theory, and therefore a natural question arises: "What would happen if we modeled the goodwill evolution using the same instruments used in the modern Finance theory?"

The first studies dealing with a stochastic extension of Nerlove and Arrow's advertising model are proposed by Tapiero [**21**], [**22**], [**23**], [**24**], [**25**]. He assumes that both the advertising policy and the forgetting phenomenon affect the goodwill evolution stochastically. He introduces transition probabilities on a discrete set of system states to model the effects of advertising on consumers' behavior and sales process and, in the feedback advertising policy framework, he obtains a stochastic differential equation (SDE for short) which models the changes of the state variable $A_t$ [**22**, p.457]. The drift term of Tapiero's SDE is similar to the Nerlove and Arrow's goodwill variation rate in (18), whereas the diffusion term is the result of some technical approximations. Other authors propose different SDEs to model the goodwill evolution, see e.g. Rao [**17**], Raman [**16**] and Sethi [**20**]. The first two authors take the r.h.s. of equation (18) as the drift term and introduce a diffusion term that focuses on a specific economic aspect of the problem (e.g. Rao introduces a white noise in the goodwill evolution).

Hence, the description of the goodwill evolution using the stochastic processes theory is not new in advertising models. However, the idea of modeling directly the goodwill evolution by controlled SDEs, as already done for the stocks evolution in Mathematical Finance, seems to be the most natural approach to the problem. The simplest stock evolution model describes the price process of a risky asset using a linear SDE. Hence, we can assume that the goodwill evolution, without any advertising flow, can be described by the following linear SDE

$$dA_t = -\delta A_t dt + \sigma_A A_t dW_t^A \ .$$

We can note that the goodwill is always positive as in the original Nerlove and Arrow's model. On the other hand the diffusion term can describe the uncertainty source due to the word-of-mouth communication, a kind of publicity which is independent of the firm advertising policy. With the above SDE we are assuming that the weight of the word-of-mouth communication is proportional to the actual goodwill and we argue that it affects the goodwill randomly. Actual consumers communicate their product experience randomly, either favorably or unfavorably. The parameter $\sigma_A \geq 0$ represents the advertising volatility and describes the power of this effect.

The control of this SDE is obtained using the advertising flow. We assume that the advertising flow increases the mean evolution of the goodwill, however the advertising message introduces also an uncertainty source in the system. These hypotheses are motivated by the assumption that the advertising policy has a double effect on the goodwill, partly deterministic (the information effect) and partly stochastic (the lure/repulsion effect). Potential consumers react randomly to advertising, being attracted or repelled by the advertising message. The advertising message has an unforeseeable effect on the goodwill, because either it may not be completely understood by the consumers, or it may completely meet the taste of the public. This can be described assuming that the goodwill evolution is represented by the following controlled SDE:

$$dA_t = (a_t - \delta A_t)\ dt + \sigma_A A_t\ dW_t^A + \sigma_a a_t\ dW_t^a \ ,$$

where $W_t^A, W_t^a$ are independent Wiener processes, while $\sigma_a$ represents the advertising volatility and describes the rate of uncertainty introduced by advertising. Under these assumptions, also the advertising flow becomes a stochastic process and now the problem can be described using the stochastic control theory

$$\begin{array}{ll} \underset{a_t \in L^2_{\mathcal{F}}(0,T;[0,+\infty))}{\text{Maximize}} & \mathbb{E}\left(\int_0^T -c\left(a_t\right)\ dt + u\left(A_T\right)\right)\ , \\ \text{Subject to} & \left\{ \begin{array}{l} dA_t = \left(a_t - \delta A_t\right)\ dt + \sigma_A A_t\, dW_t^A + \sigma_a a_t\, dW_t^a\ , \\ A_0 = A > 0\ , \end{array} \right. \end{array}$$

where $A \in \mathbb{R}$ is a constant which represents the goodwill level at the time 0.

## 3. Communication Mix

We are supposing that the goodwill is created and increased by the advertising flow, but the new marketing theories [**13**] emphasize the importance of recognizing different advertising channels which affect different segments of the market (different people react to an advertising channel in a different way). Moreover, the use of different advertising channels is necessary to meet as many consumers as possible. Finally, it is very important to characterize the most useful communication channel if the firm wants to maximize its share of the market. In order to introduce more advertising channels we can modify the goodwill motion equation as follows:

$$dA_t = \left(\sum_{i=1}^k \vartheta^i a_t^i - \delta A_t\right)\ dt + \sigma_A A_t\, dW_t^A + \sum_{i=1}^k \sigma_a^i a_t^i\, dW_t^i\ ,$$

where $\vartheta^i$ is the marginal productivity of the $i$-th advertising channel in terms of goodwill. Moreover, different advertising channels have different costs, and therefore also the objective functional has to be modified:

$$\mathbb{E}\left(-\int_0^T \sum_{i=1}^k c^i\left(a_t^i\right)\ dt + u\left(A_T\right)\right)\ ,$$

where $c^i\left(a_t^i\right)$ is cost function of the $i$-th advertising channel.
The general problem we are dealing with can be summarized as follows:

$$\tag{19} \begin{array}{ll} \underset{a_t^i \in L^2_{\mathcal{F}}(0,T;[0,+\infty))}{\text{Maximize}} & \mathbb{E}\left(-\int_0^T \sum_{i=1}^k c^i\left(a_t^i\right)\ dt + u\left(A_T\right)\right)\ , \\ \text{Subject to} & \left\{ \begin{array}{l} dA_t = \left(\sum_{i=1}^k \vartheta^i a_t^i - \delta A_t\right)\ dt \\ \qquad\qquad + \sigma_A A_t\, dW_t^A + \sum_{i=1}^k \sigma_a^i a_t^i\, dW_t^i\ , \\ A_0 = A > 0\ . \end{array} \right. \end{array}$$

The complete study of this model is very difficult because the necessary conditions for the optimality involve some forward backward SDEs that, generally, cannot be solved explicitly. However, if we consider some quadratic instances of the problem we can get some results in a closed form that clarify the features of this model. In the following Chapters we consider some instances of this problem and analyze the new effects introduced by the presence of the diffusion term in the motion equation. The following Chapters describe marketing models that can be considered as instances of (19) and make some practical and useful examples to understand the setting of the problem. Then, under the LQ hypotheses, we get some information that are used to analyze the soundness of this formulation from an economic point of view.

CHAPTER 6

# Advertising for a New Product Introduction

## 1. The problem

We present a stochastic model for the introduction of a product in the market, as a counterpart of a deterministic model proposed by Buratto and Viscolani [**5**]. We assume that the firm can control (through the advertising flow) the goodwill evolution during the programming interval $[0, T]$ and that it wants to maximize the expected utility given by the product goodwill at the (fixed) launch time $T$ and minimize the total advertising cost. The utility is a continuous and increasing function of the final goodwill level $A_T$, as in [**5**]. Here we admit an uncertainty in the goodwill evolution, which is realistic and rather obvious. A random goodwill has already been taken into account by Buratto and Viscolani [**4**] for a problem of this kind, though for the final time value only.

One of the earliest studies concerning a dynamic model in advertising is a paper written by Nerlove and Arrow [**15**]. In that work, the authors consider advertising as an investment in a stock (the goodwill $A_t$), which summarizes the effects of current and past advertising flow $a_t$. They assume that the goodwill evolution satisfies the first order linear differential equation

$$\dot{A}_t = -\delta A_t + a_t \ . \tag{20}$$

Then the goodwill decreases spontaneously with decay coefficient $\delta > 0$. $A_t$ is always positive and is sustained by the advertising investment.

The first studies dealing with a stochastic extension of Nerlove and Arrow's advertising model are proposed by Tapiero ([**21**], [**22**], [**23**], [**24**], [**25**], [**26**]). He assumes that both the advertising policy and the forgetting phenomenon affect the goodwill evolution stochastically. He introduces transition probabilities on a discrete set of system states to model the effects of advertising on consumers' behavior and sales process and, in the feedback advertising policy framework, he obtains a stochastic differential equation (SDE for short) which models the changes of the state variable $A_t$ (see [**22**, p.457]). The drift term of Tapiero's SDE is similar to the Nerlove and Arrow's goodwill variation rate in (20), whereas the diffusion term is the result of some technical approximations. Other authors propose different SDEs to model the goodwill evolution, see e.g. the paper of Rao [**17**], or the article of Raman [**16**], or the work of Sethi [**20**]. The first two authors take the r.h.s. of equation (20) as the drift term and introduce a diffusion term that focuses on a specific economic aspect of the problem (e.g. Rao introduces a white noise in the goodwill evolution).

Here we consider two uncertainty sources for the goodwill evolution and represent them by introducing two independent diffusion terms in equation (20). In fact, we may reasonably assume that both phenomena act independently on the goodwill evolution process.

The first uncertainty source is due to the word-of-mouth communication, a kind of publicity which is independent of the firm advertising policy. We assume that the weight of the word-of-mouth communication is proportional to the actual goodwill and we argue that it affects the goodwill randomly. Actual consumers communicate their product experience randomly, either favorably or unfavorably.

The second uncertainty source is due to the advertising message and is motivated by the assumption that the advertising policy has a double effect on the goodwill, partly deterministic (the information effect) and partly stochastic (the lure/repulsion effect). Potential consumers react randomly to advertising, being attracted or repelled by the product. We want to analyze in particular the optimal policies for a risk inclined decision-maker.

Let us assume that a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ satisfying the usual conditions [**28**, p.17] and two standard, stochastically independent, one-dimensional Brownian motions $W_t^A, W_t^a$ are given. Let $\{\mathcal{F}_t\}_{t\in[0,T]}$ be exactly the natural filtration generated by the Brownian motions augmented by all the $\mathbb{P}$-null sets in $\mathcal{F}$.

Moreover, given a set $U \subseteq \mathbb{R}^n$, let $L^2_{\mathcal{F}}(0,T;U)$ be the space of all the processes $X_t(\omega) : [0,T]\times\Omega \to U$ which are $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted, $U$-valued, progressively measurable, mean square integrable; i.e.

$$\mathbb{E}\left(\int_0^T \|X_t\|^2\, dt\right) < +\infty\ .$$

Finally, let $L^2_{\mathcal{F}}(\Omega; C([0,T];\mathbb{R}^n))$ be the space of all the continuous processes $X_t(\omega) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ such that $\mathbb{E}\left(\sup_{t\in[0,T]}\|X_t\|^2\right) < +\infty$.

Such technical assumptions allow us to use the stochastic optimal control theory as introduced in [**28**]; the notation used here is the same as the one defined in [**1**], [**7**], [**28**].

The problem of determining an advertising policy in order to have an optimal goodwill level at the final time $T$, when the goodwill is subject to a random evolution, is

(21)

$$\begin{array}{ll} \text{Maximize}_{a_t \geq 0} & \mathbb{E}\left(\int_0^T -c\left(a_t\right) dt + u\left(A_T\right)\right) \ , \\ \text{Subject to} & \begin{cases} dA_t = \left(a_t - \delta A_t\right) dt + \sigma_A A_t dW_t^A + \sigma_a a_t dW_t^a \ , \\ A_0 = A > 0 \ , \end{cases} \end{array}$$

where

- the function $c$ represents the cost of advertising and, according to literature (e.g. [**22**, p.455]), it is non-linear, increasing, and convex ($c' \geq 0$, $c'' \geq 0$). We assume also that $c'(0) = 0$; for the sake of simplicity, let $c'$ be a strictly increasing, unbounded, 1-to-1 map of $[0, +\infty)$ onto $[0, +\infty)$;
- the function $u$ represents the utility from the final goodwill and is homogeneous with the cost, it is twice continuously differentiable and increasing ($u' > 0$);
- $\delta > 0$ is the decay coefficient and $A > 0$ is the initial value of the goodwill;
- the first diffusion term, $\sigma_A A_t \, dW_t^A$, accounts for the word-of-mouth communication, the parameter $\sigma_A \geq 0$ is the advertising volatility;
- the second diffusion term, $\sigma_a a_t \, dW_t^a$, accounts for the lure/repulsion effect of advertising, the parameter $\sigma_a \geq 0$ is the communication effectiveness volatility.

We notice that, although the goodwill is assumed to be positive at the initial time, it may as well be negative at any time $t > 0$. We may single out three special situations which are associated with the following choices of diffusion coefficients: $\sigma_A = \sigma_a = 0$, i.e. the deterministic case with motion equation (20); $\sigma_A > 0$, $\sigma_a = 0$, i.e. the case in which the word-of-mouth phenomenon is the unique cause of randomness; $\sigma_A = 0$, $\sigma_a > 0$, i.e. the case in which the advertising lure/repulsion effect is the unique source of randomness.

## 2. General results on optimal solutions

Here we characterize the possible optimal control-state pairs by using the Peng's stochastic maximum principle [**28**]. In the following, we use the notation $[\cdot]^+$ for $\max(0, \cdot)$, as usual.

LEMMA 22. *Let $(a_t^*, A_t^*)$ be an optimal control-state pair for problem (21), and let $b$ be the inverse of the bijective map $c'$; then there exist* *processes $A_t^*, p_t, q_t^A,\ q_t^a,\ P_t, Q_t^A, Q_t^a$, which solve the following forward-backward SDE (FBSDE for short):*

$$
\left\{
\begin{array}{ll}
dA_t^* & = \left(b\left([p_t+\sigma_a q_t^a]^+\right) - \delta A_t^*\right) dt \\
 & \qquad + \sigma_A A_t^*\, dW_t^A + \sigma_a b\left([p_t+\sigma_a q_t^a]^+\right) dW_t^a\ , \\
dp_t & = \left(\delta p_t - \sigma_A q_t^A\right) dt + q_t^A\, dW_t^A + q_t^a\, dW_t^a\ , \\
dP_t & = \left(\left(2\delta - \sigma_A^2\right) P_t - 2\sigma_A Q_t^A\right) dt + Q_t^A dW_t^A + Q_t^a\, dW_t^a\ , \\
A_0^* & = A\ , \\
p_T & = u'\left(A_T^*\right)\ , \\
P_T & = u''\left(A_T^*\right)\ .
\end{array}
\right.
\tag{22}
$$

*In particular:*

$$
a_t^* \ = \ b\left([p_t+\sigma_a q_t^a]^+\right)\ . \tag{23}
$$

*Proof.* The $\mathcal{H}$-function [**28**, p.118] associated with the optimal control-state pair $(a_t^*, A_t^*)$ is

$$
\begin{aligned}
\mathcal{H}(A,a,t) = \ & (a-\delta A)\, p_t - c(a) + \tfrac{1}{2} P_t \left(\sigma_A^2 A^2 + \sigma_a^2 a^2\right) \\
& + \sigma_A A\left(q_t^A - P_t \sigma_A A_t^*\right) + \sigma_a a\left(q_t^a - P_t\sigma_a a_t^*\right)\ ,
\end{aligned}
$$

and hence

$$
\mathcal{H}_a(A,a,t) = \ p_t - c'(a) + P_t\sigma_a^2 a + \sigma_a\left(q_t^a - P_t\sigma_a a^*\right)\ . \tag{24}
$$

For all $t$, the function $\mathcal{H}_{aa}(A,a,t)$ does not depend explicitly on $A$ and the maximum condition implies that

– either $a_t^* = 0$ and $\mathcal{H}_a(A_t^*, 0, t) \le 0$, so that

$$
p_t + \sigma_a q_t^a \le c'(0) = 0\ ,
$$

– or $a_t^* > 0$ and $\mathcal{H}_a(A_t^*, a_t^*, t) = 0$, so that

$$
p_t + \sigma_a q_t^a = c'\left(a_t^*\right) > 0\ .
$$

Therefore, for all $t$, there exists a unique candidate value of the control, which can be written as $a_t^* = b\left([p_t+\sigma_a q_t^a]^+\right)$. □

In order to compute explicitly the solution of problem (21) we have to analyze the FBSDE (22). The standard method to solve this kind of equations, i.e. the so called "four step scheme" [**28**, p.387], does not work here, as the functions involved in (22) are not smooth enough. Nevertheless, we are able to propose some economic observations derived from the above Lemma.

First of all, we can obtain a representation of the solution of the adjoint equations. Let us set $\eta[a_s, b_s] = \exp\left(\int_0^t a_s ds + \int_0^t b_s dW_s\right)$, i.e.

the stochastic exponential which is the unique solution of the linear SDE $dX_t = a_t X_t dt + b_t X_t dW_t$ and let us consider the backward SDE

$$\begin{cases} dY_t &= (a_t Y_t + b_t Z_t)\, dt + Z_t dW_t \ , \\ Y_T &= \zeta \in L^2(\Omega, \mathcal{F}_T, \mathbb{P}) \ , \end{cases}$$

where $Y_t$ is a one-dimensional process, while $Z_t$ and $W_t$ are two $k$-dimensional processes. Using Itô's formula we have

$$d\left(Y_t\ \eta_t\left[-a_s, -b_s\right]\right) = (\cdots)\, dW_t \ ,$$

so that

$$\zeta\ \eta_T\left[\delta, -b_s\right] - Y_t\ \eta_t\left[-a_s, -b_s\right] = \int_t^T (\cdots)\, dW_t \ ,$$

and applying the conditional expectation $\mathbb{E}\left(\cdot|\,\mathcal{F}_t\right)$ to both sides we obtain

$$Y_t = \eta_t^{-1}\left[-a_s, -b_s\right]\ \mathbb{E}\left(\zeta\ \eta_T\left[-a_s, -b_s\right]\middle|\,\mathcal{F}_t\right) \ .$$

This well-known result [**28**, p.349] permits us to represent the processes $p_t$ and $P_t$, solution of (22), as follows:

$$\begin{aligned} p_t &= \eta_t^{-1}\left[-\delta, (\sigma_A, 0)\right]\ \mathbb{E}\left(u'\left(A_T^*\right)\ \eta_T\left[-\delta, (\sigma_A, 0)\right]\middle|\,\mathcal{F}_t\right) \ , \\ P_t &= \eta_t^{-1}\left[\sigma_A^2 - 2\delta, (2\sigma_A, 0)\right]\ \mathbb{E}\left(u''\left(A_T^*\right)\ \eta_T\left[\sigma_A^2 - 2\delta, (2\sigma_A, 0)\right]\middle|\,\mathcal{F}_t\right) \ . \end{aligned}$$

The variable $p_t$ represents the marginal value of the goodwill $A_t$ [**28**, p.251] and is always positive (as $u' > 0$). It is well-known that a deterministic optimal advertising rate satisfies the relation: "marginal cost equal to marginal goodwill value" (i.e. $c'\left(a_t^*\right) = p_t$) as documented by Sethi [**20**] and Feichtinger, Hartl and Sethi [**8**]. Here that relation holds if and only if the control has a simply deterministic effect on the goodwill evolution, i.e. if $\sigma_a = 0$. When $\sigma_a > 0$ the previous relation is modified into $c'\left(a_t^*\right) = p_t + \sigma_a q_t^a$, so that the decision-maker must also consider the uncertainty introduced by the control in the system. In fact, this is an interesting case, because, under such an assumption, the deterministic and the stochastic problems have different kinds of solution.

The variable $P_t$ represents the weight of the risk which depends on the uncertainty of the goodwill evolution related to advertising. If the decision-maker is risk averse, then $P_t < 0$ (as $u'' < 0$), so that $\mathcal{H}_{aa}(A, a, t) = -c''(a) + \sigma_a^2 P_t < 0$, for all $a$, and the function $\mathcal{H}(A, a, t)$ is concave in $a$, hence the solution (23) is indeed a maximum point of $\mathcal{H}(A, a, t)$. On the other hand, if the decision-maker is risk inclined, then $P_t > 0$ (as $u'' > 0$), so that $\mathcal{H}(A, a, t)$ may not be concave in $a$ and might also be upper unbounded w.r.t. the variable $a$ in $[0, +\infty)$. If $\sup_{a \geq 0} \mathcal{H}(A, a, t) = +\infty$, for some $t$, then the problem (21) is ill-posed, because we should have $a_t = +\infty$. This fact may occur because the variance of $A_T$ increases when the firm drives the goodwill evolution using the leverage of the process volatility. If the decision-maker is risk inclined, then the variance increment is felt as an extra-utility. In the following, we see how a well-posed deterministic problem may have a corresponding stochastic ill-posed problem, due to the risk proclivity of the decision-maker.

## 3. Deterministic interlude

If $\sigma_A = \sigma_a = 0$, then no stochastic effect is considered any more and we are faced with the following deterministic optimal control problem:

$$\begin{array}{ll} \text{Maximize}_{a_t \geq 0} & \int_0^T -c\left(a_t\right) dt + u\left(A_T\right) \ , \\ \text{Subject to} & \begin{cases} \dot{A}_t = a_t - \delta A_t \ , \\ A_0 = A > 0 \ . \end{cases} \end{array} \tag{25}$$

Here the motion equation is the same as in the Nerlove and Arrow's model, and it assures that the goodwill is always positive. The Lemma of the previous Section is specialized in the following theorem.

THEOREM 23. *Let $(a_t^*, A_t^*)$ be an optimal control-state path for problem (25), let $b$ be the inverse of the bijective map $c'$, then the optimal advertising expenditure is*

$$a_t^* = b\left(u'\left(A_T^*\right) e^{\delta(t-T)}\right) \ ,$$

*where $A_T^*$ is a solution of the equation*

$$A_T e^{\delta T} = A + \int_0^T e^{\delta s} b\left(u'\left(A_T\right) e^{\delta(s-T)}\right) ds \ . \tag{26}$$

*Proof.* If $\sigma_A = \sigma_a = 0$, then the necessary conditions (22) are satisfied by the following functions $q_t^A \equiv q_t^a \equiv Q_t^A \equiv Q_t^a \equiv 0$,

$$\begin{cases} \dot{A}_t^* &= b\left(\left[p_t\right]^+\right) - \delta A_t^* \ , \\ \dot{p}_t &= \delta p_t \ , \\ \dot{P}_t &= 2\delta P_t \ , \\ A_0^* &= A \ , \\ p_T &= u'\left(A_T^*\right) \ , \\ P_T &= u''\left(A_T^*\right) \ . \end{cases}$$

Now, the function $p_t$ can be written as $p_t = u'\left(A_T^*\right) e^{\delta(t-T)}$, so that it is always positive. Hence, the candidate optimal control is $a_t^* = b\left(u'\left(A_T^*\right) e^{\delta(t-T)}\right)$ and its associated state function can be written as

$$A_t^* = Ae^{-\delta t} + e^{-\delta t} \int_0^t e^{\delta s} b\left(u'\left(A_T^*\right) e^{\delta(s-T)}\right) ds \ . \tag{27}$$

Equation (26) is actually the transversality condition. □

**3.1. Quadratic cost and utility.** Let the cost and utility functions be as follows:

$$c\,(a) = \kappa a^2/2, \qquad u\,(A) = \gamma A^2/2, \qquad \kappa, \gamma > 0 \ ; \tag{28}$$

hence $c'(a) = ka$, $b\,(p) = p/\kappa$ and $u'\,(A) = \gamma A$. The cost function is continuous, increasing, and convex, as usual. The utility function is continuous and increasing; moreover it has been chosen quadratic and convex in order to obtain a closed form solution using the Riccati equation technique. Therefore, equation (26) is equivalent to

$$\delta\kappa A_T e^{\delta T} = \delta\kappa A - \gamma A_T \sinh\left(-\delta T\right) \ ,$$

which has the unique solution

$$A_T = \frac{\delta\kappa A}{\delta\kappa e^{\delta T} + \gamma \sinh\left(\delta T\right)} \ .$$

The above assumptions do not guarantee that the solution just obtained is optimal. In order to prove that this is indeed the optimal advertising flow we use the Riccati equation approach. As problem (25) with assumptions (28) is a linear quadratic (LQ for short) optimal control problem, we can study a Boundary Value Problem (BVP for short) in order to have necessary and sufficient conditions for optimality.

Proposition 24. *If the BVP*

$$\begin{cases} \dot{\pi}_t = 2\delta\pi_t + \pi_t^2/\kappa \ , \\ \pi_T = -\gamma \end{cases} \tag{29}$$

*has a solution defined in the whole interval* $[0, T]$*, then there exists a unique optimal control and it has the following feedback form:*

$$a_t^* = -\frac{\pi_t}{\kappa} A_t^* \ . \tag{30}$$

*Proof.* Under the assumptions (28) the optimal control problem (25) is

$$\begin{array}{ll} \text{Maximize}_{\,a_t \geq 0} & \int_0^T (-\kappa a_t^2/2)dt + \gamma A_T^2/2 \ , \\ \text{Subject to} & \begin{cases} \dot{A}_t = a_t - \delta A_t \ , \\ A_0 = A > 0 \ . \end{cases} \end{array} \tag{31}$$

The feasible set of the problem is $L^2\left(0, T; [0, +\infty)\right)$. The relaxed problem, with the feasible set $L^2\left(0, T; \mathbb{R}\right)$, is an LQ optimal control problem.

We prove that an optimal control $a_t^* \in L^2(0, T; \mathbb{R})$ of the relaxed problem must satisfy the constraint $a_t^* \geq 0$ and must be an optimal control of the original problem (31).

Using a well-known result on the LQ deterministic optimal control problems [**28**, p.296] we know that if a solution $\pi_t$ of the BVP (29) is given , then the relaxed LQ optimal control problem has an optimal control which can be written in the feedback form (30). Now, the ODE in (29) has two equilibrium points: $\pi^e = 0$, which is unstable, and $\pi^\varepsilon = -2\delta\kappa < 0$, which is asymptotically stable. The condition $\pi_T = -\gamma$ can only be satisfied if the BVP solution is negative in the whole interval $[0, T]$. Therefore we have that $a_t^* = -\pi_t A_t^*/\kappa > 0$ for all $t$ and the optimal control of the relaxed problem is feasible (and optimal) also for the original problem (31). $\square$

The BVP (29) can be solved locally in $(T - \varepsilon, T]$ for some $\varepsilon > 0$, and its explicit solution is

$$\pi(t) = -\frac{\delta\kappa\gamma e^{\delta(t-T)}}{\delta\kappa e^{\delta(T-t)} + \gamma \sinh(\delta(T-t))} \,. \tag{32}$$

We observe that the denominator $\delta\kappa e^{\delta(T-t)} + \gamma \sinh(\delta(T-t))$ is strictly decreasing in $[0, T]$ and reaches the value $\delta\kappa > 0$ at $t = T$, so that it is strictly positive for all $t \in [0, T]$. Therefore function (32) solves the BVP (29) for all $t \in [0, T]$.

In order to study the feedback function (30) it is convenient to analyze the phase portrait (Figure 6.1), rather than using the explicit solution (32).

We observe that the stable equilibrium point $\pi^\varepsilon$ is an attractor for the set $(-\infty, 0)$. On the other hand, all the solutions starting from a point of the set $(0, +\infty)$ go to $+\infty$ as $t \mapsto +\infty$. To respect the condition $\pi_T = -\gamma$, we have that either $\lim_{t \mapsto -\infty} \pi_t = 0$ (when $-\gamma \in (-2\delta\kappa, 0)$, i.e. $\gamma/2\kappa < \delta$), or $\lim_{t \mapsto -\infty} \pi_t = -\infty$ (when $-\gamma \in (-\infty, -2\delta\kappa)$, i.e. $\gamma/2\kappa > \delta$). Then we are led to consider three different kinds of policies (Figure 6.2).

- **$\gamma/2\kappa < \delta$:** the decay coefficient $\delta$ is large, therefore the optimal advertising/goodwill ratio (i.e. $-\pi_t/\kappa$) is strictly increasing and upper bounded by the value $\gamma/\kappa$;
- **$\gamma/2\kappa = \delta$:** the optimal advertising/goodwill ratio is constant during the whole programming interval;
- **$\gamma/\kappa 2 > \delta$:** the decay coefficient $\delta$ is small, therefore the optimal advertising/goodwill ratio is strictly decreasing and is lower bounded by the value $\gamma/\kappa$.

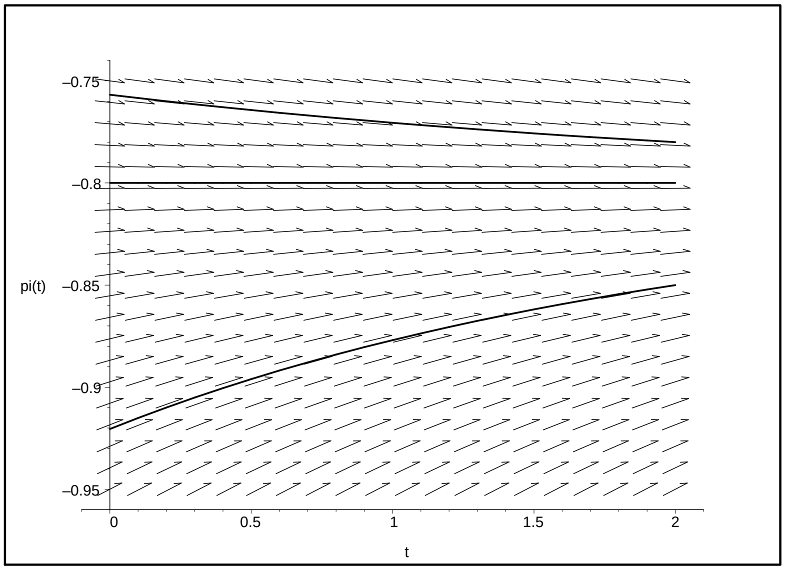


FIGURE 1. Trajectories of $\pi_t$

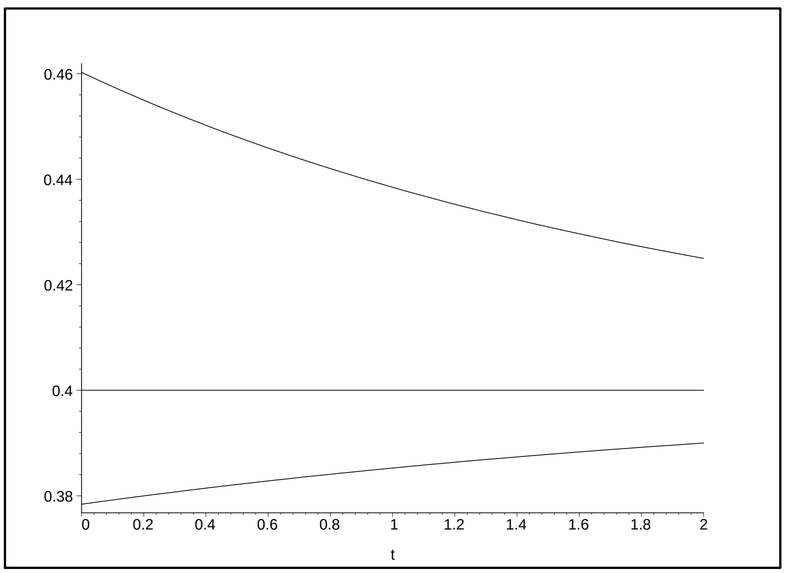


FIGURE 2. Advertising/goodwill ratio trajectories

## 4. Stochastic framework and risk proclivity

In the previous Section we have seen how the LQ instance of problem (25) can be profitably used in order to obtain some explicit results. Here we focus on the analogous instance of problem (21), where the cost and utility functions are $c(a) = \kappa a^2/2$ and $u(A) = \gamma A^2/2$, with

$\kappa, \gamma > 0$ fixed parameters. In the stochastic framework the convexity of the utility function has a precise interpretation: it means assuming a risk inclined decision-maker. With such a choice we could describe a firm which wants to obtain a share of a new market and is prepared to accept some risks in order to reach that goal.
The stochastic problem (21) with the quadratic cost and utility functions is
(33)

$$\begin{array}{ll} \text{Maximize}_{a_t \geq 0} & \mathbb{E}\left(\int_0^T (-\kappa a_t^2/2)dt + \gamma A_T^2/2\right) \ , \\ \text{Subject to} & \begin{cases} dA_t = (a_t - \delta A_t)\, dt + \sigma_A A_t dW_t^A + \sigma_a a_t dW_t^a \ , \\ A_0 = A > 0 \ . \end{cases} \end{array}$$

As in the deterministic case, the problem is a linear quadratic (stochastic) optimal control problem as soon as we relax the constraint $a_t \in L^2_{\mathcal{F}}(0, T; [0, +\infty))$. Now, we show that if a BVP called stochastic Riccati equation has a solution, then the relaxed problem has an optimal solution, which is related to the solution of the BVP. Moreover, the control is non-negative, so that the solution is optimal for the original problem, too.

THEOREM 25. *If the BVP*

$$\begin{cases} \dot{\pi}_t = (2\delta - \sigma_A^2)\, \pi_t + \pi_t^2 / \left(\kappa + \sigma_a^2 \pi_t\right) \ , \\ \pi_T = -\gamma \ , \\ \kappa + \sigma_a^2 \pi_t > 0 \end{cases} \tag{34}$$

*has a solution which is defined in the whole programming interval* $[0, T]$*, then there exists a unique optimal control for problem (33), which has the feedback form:*

$$a_t = -\frac{\pi_t}{\kappa + \sigma_a^2 \pi_t} A_t \ . \tag{35}$$

*Proof.* Let us consider the relaxed instance of the problem (33) (i.e. the problem with the feasible set $L^2_{\mathcal{F}}(0, T; \mathbb{R})$). Using the solution of the BVP (34) we can write the objective functional of (33) as (see [**28**, p.315]:

$$\mathbb{E}\left(\cdots\right) = -\frac{1}{2}\pi_0 A^2 - \frac{1}{2}\mathbb{E}\left(\int_0^T \sqrt{\kappa + \sigma_a^2 \pi_t}\ \left| a_t + \frac{\pi_t}{\kappa + \sigma_a^2 \pi_t} A_t \right| dt\right) \ .$$

Therefore the optimal control is unique and has the feedback form (35). Moreover, if we choose such a control, the motion equation becomes a linear SDE with a stochastic exponential solution. This guarantees that the optimal state process is positive, $A_t^* > 0$. In order to state that also the optimal control is positive we have to prove that the solution of equation (34) is negative. This depends on the fact that the ODE in (34) has $\pi = 0$ as an equilibrium point and moreover $\pi_T = -\gamma < 0$. $\square$

We observe that, as $\kappa > 0$ the BVP (34) is solvable if and only if the problem (34) is well-posed [**7**], [**1**].

**4.1. Word-of-mouth effect only.** If $\sigma_a = 0$ and $\sigma_A > 0$, then only the word-of-mouth effect is considered in the model. The BVP (34) becomes:

$$\begin{cases} \dot{\pi}_t = (2\delta - \sigma_A^2)\,\pi_t + \pi_t^2/\kappa \ , \\ \pi_T = -\gamma \ , \\ \kappa > 0 \ . \end{cases} \tag{36}$$

First, we notice that the third condition is satisfied trivially. Moreover, the above BVP is essentially the same as (29) in the deterministic instance. They only differ for the linear term coefficient, which now is $2\delta - \sigma_A^2$ and may be negative as well as positive. Then we have to consider the following two different cases.

$2\delta - \sigma_A^2 > 0$**:** the ODE in (36) has the same phase portrait as the ODE in (29). The two equilibrium points are now $\pi^e = 0$, and $\pi^\varepsilon = -(2\delta - \sigma_A^2)\,\kappa < 0$. If $\sigma_a = 0$ and $2\delta > \sigma_A^2 > 0$, then the stochastic optimal control problem (33) is equivalent to (i.e. has the same feedback form solution as) the following deterministic optimal control problem:

$$\begin{array}{ll} \text{Maximize}_{\,a_t \geq 0} & \int_0^T (-a_t^2\kappa/2)dt + \gamma A_T^2/2 \ , \\ \text{Subject to} & \begin{cases} dA_t = (a_t - (2\delta - \sigma_A^2)\,A_t/2)\,dt \ , \\ A_0 = A > 0 \ . \end{cases} \end{array}$$

This implies that there exists a trade-off between the word-of-mouth publicity and the spontaneous decay of the goodwill due to the consumers forgetting phenomenon.

$2\delta - \sigma_A^2 \leq 0$**:** the phase portrait of the ODE in (29) is different from the previous one because both equilibrium points are in the non-negative half-plane. $\pi^e = 0$ becomes an attractor for the set $(-\infty, 0)$, therefore the state coefficient in the feedback function is strictly decreasing and is lower bounded by the value $\gamma$. Also in the first part of the programming interval a strong advertisement is convenient because the goodwill is self-supporting (now the goodwill increases spontaneously, hence it is convenient to increase the process immediately).

**4.2. Lure/repulsion effect only.** If $\sigma_A = 0$ and $\sigma_a > 0$, then only the lure/repulsion effect is considered in the model. The Riccati BVP (34) becomes:

$$\begin{cases} \dot{\pi}_t = 2\delta\pi_t + \pi_t^2 / \left(\kappa + \sigma_a^2 \pi_t\right) \ , \\ \pi_T = -\gamma \ , \\ \pi_t \sigma_a^2 + \kappa > 0 \ . \end{cases} \tag{37}$$

Here, the third condition is essential in order to obtain a well-posed problem. The ODE in (37) has two equilibrium points: $\pi^e = 0$, an unstable point, and $\pi^\varepsilon = -2\delta\kappa / \left(1 + 2\delta\sigma_a^2\right) < 0$, which is asymptotically stable. The latter, $\pi^\varepsilon$, is an attractor for the set $(-\infty, 0)$. On the other hand, all the solutions starting from points of the set $(0, +\infty)$ go to $+\infty$ as $t \mapsto +\infty$.

The following cases may occur:

**$\pi^\varepsilon < -\gamma$:** necessarily $\kappa - \gamma\sigma_a^2 > 0$ and the decay coefficient $\delta$ is large, $\delta > \gamma/2(\kappa - \gamma\sigma_a^2)$. The solution $\pi_t$ is negative, strictly decreasing and lower bounded by the value $-\gamma$. From $\pi_t \geq -\gamma$ we obtain that $\pi_t\sigma_a^2 + \kappa \geq k/(1 + 2\delta\sigma_a^2) > 0$, so that the third condition in (37) is satisfied, too. The advertising cost (which depends on $k$) is high enough so that the problem is well-posed. The state coefficient in the feedback function can be rewritten as

$$-\frac{\pi_t}{\kappa + \sigma_a^2 \pi_t} = \frac{1}{\sigma_a^2}\left(\frac{\kappa}{\kappa + \sigma_a^2 \pi_t} - 1\right) \ ,$$

so that we observe that the optimal advertising/goodwill ratio is strictly increasing and upper bounded by $\gamma / \left(\kappa - \sigma_a^2\gamma\right)$.

**$\pi^\varepsilon = -\gamma$:** as previously $\kappa - \gamma\sigma_a^2 > 0$ (the advertising cost must be relatively high) and the decay coefficient is $\delta = \gamma/2(\kappa - \gamma\sigma_a^2)$. As in the first case, the third condition in (37) is satisfied. The solution $\pi_t$ and the optimal advertising/goodwill ratio are constant.

**$\pi^\varepsilon > -\gamma$:** if $\kappa - \gamma\sigma_a^2 > 0$ (relatively high advertising cost), then the decay coefficient is small, $\delta < \gamma/2(\kappa - \gamma\sigma_a^2)$, otherwise, if $\kappa - \gamma\sigma_a^2 \leq 0$ (relatively low advertising cost), the decay coefficient is not subject to any additional constraints, $\delta > 0$. In either alternative, the solution $\pi_t$ of the ODE in (37) is negative, strictly increasing and upper bounded by the value $-\gamma$. Moreover, $\lim_{t \mapsto -\infty} \pi_t = -\infty$, therefore we can define $t^* \in \mathbb{R}$ such that

$$t^* = \inf\left\{t \in \mathbb{R} : \kappa + \sigma_a^2 \pi_t \geq 0\right\} \ .$$

Now, the third condition in (37) is satisfied if and only if $t^* < 0$. In that case, the optimal advertising/goodwill ratio is strictly decreasing and reaches its minimum $\gamma / \left(\kappa - \sigma_a^2 \gamma\right)$ at $t = T$.

## 5. Conclusion

From the comparison of the deterministic and the stochastic LQ instances of the problem (21) we have found some new features in the stochastic viewpoint, which are interesting from an economic point of view.

Under the LQ assumption the decision-maker is risk inclined, therefore when she/he maximizes a quadratic final utility she/he is choosing an advertising policy that increases as much as possible both the mean and the variance of the final goodwill random variable $A_T$.

The presence of the word-of-mouth effect compensate partly the goodwill spontaneous decay. This phenomenon can also reverse the goodwill decay when the goodwill volatility is sufficiently high. However, the presence of this random advertising channel cannot modify deeply the feedback form of the optimal advertising policy. It can shift the equilibrium point $\pi^{\varepsilon}$ to the positive half-plane, and so doing reduce the possibility to have an increasing optimal advertising/goodwill ratio. Nevertheless, it does not modify the feedback form, nor transform a deterministic well-posed problem into a stochastic ill-posed one.

On the other hand, the action of the uncertainty connected with the lure/repulsion effect of the advertising message is deeper (and it is not so surprising because the advertising policy affects *directly* the evolution of the goodwill variance). This phenomenon introduces a high uncertainty in the system, so that the stochastic analogous of a well-posed deterministic problem may as well be an ill-posed stochastic problem for a risk inclined decision-maker. When the stochastic problem is well-posed, the optimal advertising/goodwill ratio has a form which differs from the one of the analogous deterministic problem. The new form shifts the equilibrium point $\pi^{\varepsilon}$ to the positive half-plane and reduces, as observed above, the possibility to have an increasing optimal advertising/goodwill ratio.

From a marketing point of view, these results suggest two kinds of advertising campaign for a risk inclined decision-maker: assuming that the mean goodwill increases, the first kind of advertising campaign is rather prudent and concentrates the advertising flow near the end of the programming interval. The second kind is riskier and drives the goodwill quickly to a high level in order to exploit the possible positive fluctuations of the state variable. In order to select one of the two alternatives a decision-maker has to consider the trade-off between the deterministic effects of the advertising channel (as the advertising cost or the goodwill decay) and the high promising (but risky) effects connected to the stochastic part of the advertising message.

CHAPTER 7

# Event Planning, Communication Mix, and Congestion

## 1. Model description

We consider an organization which has planned a social event (concert, workshop, soccer match), at a fixed time $T$, in a place with a limited number of seats. An advertising campaign has to be organized in order to stimulate as wide a participation as possible. We suppose that the customers pay the tickets only at time $T$ and that the demand depends on the event goodwill at $T$. Moreover, we assume that the goodwill is created and increased by the advertising for the event, possibly by using different channels (e.g. TV, radio, internet, newspapers, magazines). We consider two different kinds of advertising channels. The first one affects the goodwill evolution in a deterministic way (which we call the deterministic channel): the advertising message is correctly understood by the consumers and contributes to increase the goodwill value or to slow down a goodwill decrease. The second one (which we call the stochastic channel) has a double effect: it surely increases, or slows down a decrease of the goodwill expectation, but it also introduces some uncertainty in the goodwill evolution. The stochastic channel advertising message has an unforeseeable effect on the goodwill, because either it may not be completely understood by the consumers, or on the contrary it may meet the taste of the public very well. In this paper, we initially assume that the organization can use only one advertising channel, then we analyze the case in which two different advertising channels are allowed. We want to compare the different optimal policies obtained when the channels are deterministic or stochastic, in order to understand how the presence of some uncertainty modifies the effectiveness of the advertising message.

Furthermore, we assume that a part of the total demand may not be satisfied because of a congestion phenomenon due to a limited number of seats. Therefore a well-planned communication program should assure that the revenue from the sold tickets fully cover all the costs. One of the organization's objectives is to control the goodwill evolution in order to obtain a demand as close as possible to the congestion threshold. A lower level means that some tickets have not been sold, so that the revenue is less than the potential maximum one. On the contrary a higher level would signify that the demand is greater than the offer; in such a way there exist some unsatisfied customers who cause an unfavorable word-of-mouth publicity, with a negative effect on the success of future events. Moreover, the organization must take into account the consumers' satisfaction in order to obtain a well-organized event. Another objective we consider is the minimization of the total advertising cost, which is typical in the firm theory (see. [**27**, cap. I]).

We refer to the concept of goodwill, introduced by Nerlove and Arrow in [**15**], which resumes the effects of advertising on the demand and whose evolution is described by a linear differential equation. We consider two control functions, $a_t$ and $v_t$, which represent the advertising flows for the two channels, with marginal productivity $\vartheta > 0$ and $\rho > 0$, respectively, in terms of goodwill. We assume that the goodwill is a stochastic process driven by those two different controls. Let $v_t$ be the control function which deterministically affects the goodwill evolution, i.e. the advertising flow that only increases the mean evolution. On the other hand, let $a_t$ be the control function associated with the stochastic advertising channel, that increases both the mean and the variance of the stochastic process.

Let us denote by $A_t$ the goodwill level at time $t \in [0, T]$. We assume that the organization image, as well as the event features, contribute to affect the initial goodwill of the event, $A_0$, which is therefore positive: i.e. $A_0 = \underline{A} > 0$. Assuming that the publicity campaign is short enough, we can suppose that the consumers do not forget the advertising messages, therefore the goodwill decay is negligible. Hence, the goodwill decay term in the Nerlove and Arrow's model is assumed to be zero. Moreover, we assume that the goodwill motion equation is

$$dA_t = (\vartheta a_t + \rho v_t)\ dt + \sigma a_t\ dW_t\ ,$$

where $\sigma \geq 0$ is the stochastic channel uncertainty coefficient and $W_t$ is a standard Brownian motion. In other words, the goodwill is the solution of a controlled linear stochastic differential equation.

The stochastic approach to the communication mix problem is rather innovative. The integration of two communication forms has been studied under deterministic assumptions in [**2**] and [**3**]. Whereas, the problem of planning a publicity campaign for an event that takes place at a fixed time has some analogies with the problem of introducing a product in the market, which has already been studied from a deterministic point of view in [**4**]. At a more general and different level, there are a lot of papers concerning stochastic models in advertising:

[**16**], [**17**], [**20**], [**21**], [**22**], [**23**], [**24**], [**25**], [**26**]. For a complete survey of dynamic optimal control models in advertising, updated to the year 1994, we refer to [**8**].

As far as the organization's objectives are concerned, we assume that the advertising costs have a quadratic form. In particular, after denoting by $\kappa > 0$ and $\beta > 0$ the cost coefficients of the two advertising channels, the total advertising cost is assumed to be $\int_0^T \left(\kappa a_t^2 + \beta v_t^2\right) dt$ (which is the same quadratic cost assumption for an advertising channel as used in [**10**]).

Let us denote by $d$ the demand function, $d : [0, +\infty) \to [0, +\infty)$, which we assume to be an increasing and bijective function of the final goodwill level $A_T$. Let $\bar{d} > 0$ be the congestion threshold, e.g. the maximum number of available seats, then there exists a corresponding goodwill congestion threshold $\bar{A}$, such that the demand fully covers the offer, i.e. $d(\bar{A}) = \bar{d}$. The sales are given by the following function of the demand:

$$s\left(A_T\right) = \min\left\{d\left(A_T\right), \bar{d}\right\} \;=\; d\left(\min\left\{A_T, \bar{A}\right)\right\} \;.$$

The organization wants to achieve a final goodwill level $A_T$ as close as possible to the goodwill congestion threshold $\bar{A} > \underline{A}$. In order to reach such a result, the firm defines a loss function $\ell$ of the final goodwill level $A_T$ with the properties that $\ell\left(A_T\right) > 0$ and $\ell\left(A_T\right) = 0$ if and only if $A_T = \bar{A}$. In this paper we assume that the loss function chosen by the firm is quadratic, this is in fact a typical choice in order to obtain a computational tractable problem [**19**]. Let $\gamma\left(A_T - \bar{A}\right)^2$ be the loss function the organization wants to minimize, where $\gamma > 0$ is a penalty coefficient, which represents the cost for the distance of the target value from the final goodwill value.

The paper is organized as follows: in Section 2 the problem is analyzed from a mathematical point of view; the optimal policy in a feedback form is found using the Riccati equation technique. In Section 3 we comment on this solution assuming that the organization can use only one advertising channel, either deterministic, or stochastic. Then, in Section 4, we deal with the general solution and compare the efficiency and the risk of the two channels.

## 2. The problem

Let us assume that a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ satisfying the usual conditions [**28**, p.17] and a standard Brownian motion $W_t$ are given. $\{\mathcal{F}_t\}_{t\in[0,T]}$ is the natural filtration generated by $W_t$ augmented by all the $\mathbb{P}$-null sets in $\mathcal{F}$. Such technical assumptions allow us to use the stochastic optimal control theory [**28**]. The model described in the previous Section can be written as the following optimal control problem

(38)
$$\begin{array}{ll} \underset{a_t, v_t \geq 0}{\text{Minimize}} & \overline{J}[a_t, v_t] = \frac{1}{2}\mathbb{E}\left(\int_0^T \left(\kappa a_t^2 + \beta v_t^2\right) dt + \gamma\left(A_T - \bar{A}\right)^2\right) \ , \\ \text{Subject to} & \left\{ \begin{array}{l} dA_t = \left(\vartheta a_t + \rho v_t\right) \ dt + \sigma a_t \ dW_t \ , \\ A_0 = \underline{A} > 0 \ , \end{array} \right. \end{array}$$

where the control functions are non-negative, $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted, $\mathbb{R}$-valued, square-integrable processes defined on the time interval $[0,T]$: $a_t, v_t \in L^2_{\mathcal{F}}\left(0, T; [0, +\infty)\right)$.
Moreover, let $L^2_{\mathcal{F}}\left(\Omega; C\left([0,T]; \mathbb{R}\right)\right)$ be the set of $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted, $\mathbb{R}$ valued, continuous processes $Y_t$ defined on the time interval $[0,T]$ and such that $\mathbb{E}\left(\sup_{t\in[0,T]}|Y_t|^2\right) < +\infty$ , then the state function $A_t$ is the unique process in $L^2_{\mathcal{F}}\left(\Omega; C\left([0,T]; \mathbb{R}\right)\right)$ that solves the motion equation in (38).

In order to deal with a standard LQ stochastic control problem, it is convenient to introduce a goodwill translation and to relax the constraint $a_t, v_t \geq 0$. Let us define the new state variable (translated goodwill) $X_t = A_t - \bar{A}$. After setting $u_t = (a_t, v_t)' \in L^2_{\mathcal{F}}\left(0, T; \mathbb{R}^2\right)$ and

$$R = \begin{pmatrix} \kappa & 0 \\ 0 & \beta \end{pmatrix}, \qquad B = \begin{pmatrix} \vartheta & ,\rho \end{pmatrix}, \qquad D = \begin{pmatrix} \sigma & ,0 \end{pmatrix},$$

the following optimal control problem is the relaxed form of problem (38) (as usual the transpose of a matrix $M$ is written as $M'$):

$$\begin{array}{ll} \underset{u_t \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^2)}{\text{Minimize}} & J[u_t] = \frac{1}{2}\mathbb{E}\left(\int_0^T u_t' R u_t dt + \gamma X_T^2\right) \ , \\ \text{Subject to} & \left\{ \begin{array}{l} dX_t = B u_t \ dt + D u_t \ dW_t \ , \\ X_0 = X = \underline{A} - \bar{A} < 0 \ . \end{array} \right. \end{array} \tag{39}$$

We solve this problem using the *completion of squares* method [**28**] which leads to the following proposition.

PROPOSITION 26. *Let us assume that there exists* $\pi_t \in C^1\left([0,T]; \mathbb{R}\right)$ *solution of the Riccati BVP*

$$\left\{ \begin{array}{l} \dot{\pi}_t = \pi_t^2 B \left(R + D'\pi_t D\right)^{-1} B' \ , \\ \pi_T = \gamma \ , \end{array} \right. \tag{40}$$

*such that the matrix* $R + D'\pi_t D$ *is positive definite (i.e.* $R + D'\pi_t D \succ 0$*), then there exists a unique optimal control* $u_t^*$ *for the problem (39) and it can be written in the following feedback form:*

$$u_t^* = -\left(R + D'\pi_t D\right)^{-1} B' \pi_t X_t \ . \tag{41}$$

*Proof.* This Proposition is a standard result called stochastic Riccati equation [**28**, p.315, Th.6.1]. For the reader's convenience we sketch the proof under the previous assumptions.

Let us consider the process $\pi_t X_t^2$, where $X_t$ is the solution of the motion equation in (39) for some control process $u_t$. The stochastic differential of this process is:

$$\begin{aligned} d\left[\pi_t X_t^2\right] = &\dot{\pi}_t X_t^2 dt + 2\pi_t X_t B u_t dt \\ &+ 2\pi_t X_t D u_t dW_t + u_t' D' \pi_t D u_t dt \ , \end{aligned}$$

hence, as $\pi_T = \gamma$ and $X_0 = X$, we obtain

$$\mathbb{E}\left(\gamma X_T^2\right) = \pi_0 X^2 + \mathbb{E}\left(\int_0^T \dot{\pi}_t X_t^2 + 2\pi_t X_t B u_t + u_t' D' \pi_t D u_t dt\right) \ .$$

By substituting $\dot{\pi}_t$ in the previous equation and setting $\hat{R}_t = R + D'\pi_t D \succ 0$, we have

$$\begin{aligned} &\mathbb{E}\left(\textstyle\int_0^T u_t' R u_t dt + \gamma X_T^2\right) = \\ &= \pi_0 X^2 + \mathbb{E}\left(\textstyle\int_0^T B\hat{R}_t^{-1} B' \left(\pi_t X_t\right)^2 + 2\pi_t X_t B u_t + u_t' \hat{R}_t u_t dt\right) \ . \end{aligned}$$

From the assumption $\hat{R}_t \succ 0$ it follows that there exists $\tilde{R}_t$, a $2 \times 2$ positive definite symmetric matrix, such that $\tilde{R}_t \tilde{R}_t = \hat{R}_t$. Hence, we have that

$$\mathbb{E}\left(\int_0^T u_t^T R u_t dt + \gamma X_T^2\right) = \pi_0 X^2 + \left\|\tilde{R}_t^{-1} B' \pi_t X_t + \tilde{R}_t u_t\right\|_{L^2}^2 \ .$$

The previous relation holds for all $u_t \in L^2_{\mathcal{F}}\left(0, T; \mathbb{R}\right)$ , therefore a minimum point is obtained if and only if we choose $u_t^*$ such that

$$\left\|\tilde{R}_t^{-1} B' \pi_t X_t + \tilde{R}_t u_t\right\|_{L^2}^2 = 0 \ ,$$

which implies that $\tilde{R}_t^{-1} B' \pi_t X_t + \tilde{R}_t u_t = 0$, i.e. $u_t = -\hat{R}_t^{-1} B' \pi_t X_t$. $\square$

The optimal solution of problem (39) is connected to the solution of BVP (40); moreover, the optimal value of the objective functional depends on the initial value of the function $\pi_t$. If $(u_t^*, X_t^*)$ is the optimal control-state pair of (39), then the following equation holds:

$$\mathbb{E}\left(\frac{1}{2}\int_0^T \left(u_t^*\right)^T R u_t^* dt + \frac{1}{2}\gamma \left(X_T^*\right)^2\right) = \frac{1}{2}\pi_0 X^2 \ . \tag{42}$$

We note that $R + D^T \gamma D \succ 0$, therefore in the interval $[s, T]$ (with $s < T$) there exists a solution of (40) such that $R + D^T \pi_t D \succ 0$. In the following proposition we prove that this solution can be extended to the whole programming interval $[0,T]$; hence we prove the existence of the optimal control for problem (39).

PROPOSITION 27. *There exists a positive and strictly increasing function $\pi_t \in C^1([0,T];\mathbb{R})$ that solves (40).*

*Proof.* The BVP (40) can be rewritten as

$$\begin{cases} \dot{\pi}_t = \pi_t^2 \left(\beta\vartheta^2 + \kappa\rho^2 + \pi_t\rho^2\sigma^2\right) / \left(\kappa + \pi_t\sigma^2\right)\beta \ , \\ \pi_T = \gamma \ . \end{cases}$$

Now we observe that $\pi_t = 0$ is the unique equilibrium point in the positive half-plane and that it is unstable. Moreover, all the solutions starting with a positive initial value are strictly increasing and go to $+\infty$ as $t \mapsto +\infty$. A solution of (40) cannot cross an equilibrium point; therefore, in order to have $\pi_T = \gamma$, the solution of (40) must start at a positive value (i.e. $\pi_0 > 0$) and must be a strictly increasing function. $\square$

The previous statements hold under the assumption that the matrix $R + D^T\pi_t D$ is positive definite. In fact a simple computation shows that, as $\pi_t > 0$, the matrix $R + D^T\pi_t D$ is positive definite, hence problem (39) has a unique solution, which is given by the feedback form (41).

The following Theorem closes this section showing that the relaxed problem provides exactly the optimal solution of the original one.

THEOREM 28. *If $(u_t^*, X_t^*)$ is an optimal solution of problem (39), and the control is $u_t^* = (a_t^*, v_t^*)'$, then $((a_t^*, v_t^*), A_t^*)'$, with $A_t^* = X_t^* + \overline{A}$, is an optimal state-control path for problem (38).*

*Proof.* We note that the feasible set $L^2_{\mathcal{F}}(0,T;[0,+\infty))$ of problem (38) is a subset of the feasible set of problem (39), moreover the two optimal control problems have objective functionals with the same structure. The following relation holds: if a given control $u_t = (a_t, v_t)'$ belongs to $L^2_{\mathcal{F}}(0,T;[0,+\infty))$, then $\overline{J}[a_t, v_t] = J[u_t]$, hence, if an optimal control function for (39) is also feasible for (38), it is an optimal control for (38), too.

Let $X_t^*$ be the optimal state process of problem (39), then it is the solution of the following linear SDE

$$\begin{cases} dX_t^* = -B\left(R + D'\pi_t D\right)^{-1} B'\pi_t X_t^* dt \\ \qquad\qquad -D\left(R + D'\pi_t D\right)^{-1} B'\pi_t X_t^* dW_t \ , \\ X_0^* = X < 0 \ . \end{cases}$$

This solution can be represented using the stochastic exponential process

$$\eta_t \left[a_s, b_s\right] \triangleq \exp\left(\int_0^t \left(a_s - \frac{1}{2}b_s^2\right) ds + \int_0^t b_s dW_s\right)$$

in the following form:

$$X_t^* = X_0^* \ \eta_t \left[-\left(B\left(R + D'\pi_s D\right)^{-1} B'\pi_s\right), -D\left(R + D'\pi_s D\right)^{-1} B'\pi_s\right],$$

where the terms inside the square brackets are negative and therefore the process $X_t^*$ is also negative for all $t \in [0, T]$. We observe that the optimal control pair $(a_t^*, v_t^*)$ of problem (38) is given in a feedback form by equation (41). Therefore the components $a_t^*$ and $v_t^*$ are both positive for all $t \in [0, T]$ and hence they are feasible for problem (38), too. We can conclude that, as the optimal solution $(u_t^*, X_t^*)$ for problem (39) is feasible for problem (38), it is also optimal for problem (38). □

COROLLARY 29. *Under the general assumption of this Section, problem (38) has an optimal solution which is uniquely determined solving the relaxed problem (39).*

*Proof.* The relaxed problem (39) has an optimal solution because the solution of the Riccati equation is defined in the whole programming interval [**1**], [**7**]. Hence, we can conclude the proof using the previous result. □

We note that the left-hand side of equation (42) gives the sum of the cost for the advertising campaign and of the total penalty to pay. The whole cost is proportional to the square of the difference between the initial goodwill value $\underline{A}$ and the target goodwill value $\bar{A}$, with proportionality factor $\pi_0$.

## 3. One advertising channel

In the following we tackle the problem characterized by only one advertising channel and we distinguish the case in which such a channel is deterministic from the case in which it is stochastic.

**3.1. Deterministic channel.** Let us assume that the decision-maker can use only one deterministic advertising channel. This is the case of problem (38) with $\vartheta = \sigma = 0$. Under these assumptions we obtain a deterministic optimal control problem. The optimal feedback control (41) is

$$v_t^* = w_t \left(\bar{A} - A_t^*\right) \ , \tag{43}$$

where $w_t = (\rho\pi_t)/\beta$ (we say that $w_t$ is the *weight* of $v_t$ in the feedback function), and $\pi_t$ is the solution of the following BVP,

$$\begin{cases} \dot{\pi}_t = \rho^2\pi_t^2/\beta \ , \\ \pi_T = \gamma \ . \end{cases} \tag{44}$$

We observe that the weight of $v_t$ is proportional to the ratio between the marginal productivity of the advertising channel and its cost coefficient. As $\pi > 0$, we have that the greater the marginal productivity is, the greater the advertising effect becomes. On the contrary, the greater the cost coefficient is, the smaller the advertising effect becomes.

**3.2. Stochastic channel.** Let us assume now that the organizer can use only the stochastic advertising channel (i.e. $\rho = 0$); then problem (38) has the following optimal feedback control (41)

$$a_t^* = m_t\left(\bar{A} - A_t^*\right) \ ,$$

where

$$m_t = \frac{\vartheta\pi_t}{\kappa + \sigma^2\pi_t}$$

is the *weight* of $a_t$ in the feedback function. The Riccati equation (40) becomes

$$\begin{cases} \dot{\pi}_t = \vartheta^2\pi_t^2/\left(\kappa + \sigma^2\pi_t\right) \ , \\ \pi_T = \gamma \ . \end{cases}$$

**3.3. Different behaviors.** In order to analyze the difference between the two behaviors in the one-channel instances of the problem, we compare the two weight functions $w_t$ and $m_t$ when $\vartheta = \rho$ (i.e. the two coefficients representing the marginal productivity are the same) and $\kappa = \beta$ (i.e. the marginal cost of the two advertising channels is the same). First, we solve the problem assuming that the control has a deterministic effect on the goodwill evolution, then we tackle the problem with the same parameters values, but under the assumption that the control affects the goodwill evolution stochastically.
We observe that both weight functions $w_t$ and $m_t$ are increasing, and $w_t$ is greater than $m_t$ in a left neighborhood of $T$. This means that the deterministic channel is more intensively used at the end of the campaign, because the effect of its control on the state is exactly known, whereas the stochastic channel has to be used more prudently, because the risk represents an extra cost to take into account.

## 4. Two advertising channels

In the following we tackle the problem characterized by two advertising channels; we first consider the case in which they are both deterministic and then we analyze the case in which one of them is stochastic. A more realistic situation considers both advertising channels as stochastic, since generally their effects on the goodwill are not completely predictable. However, it is reasonable to assume that the two channels affect the goodwill evolution in different ways and that one of them has a behavior closer to the deterministic channel.

**4.1. Deterministic channels.** Let us assume that the organizer can use both channels and that the two control functions have only a deterministic effect on the goodwill evolution (i.e. we set $\sigma = 0$ in problem (38)). We obtain a deterministic optimal control problem and the optimal feedback control (41) is

$$\begin{pmatrix} a_t^* \\ v_t^* \end{pmatrix} = \begin{pmatrix} m_t \\ w_t \end{pmatrix} \left( \bar{A} - A_t^* \right) \ ,$$

where $m_t = (\pi_t \vartheta) / \kappa$ , $w_t = (\pi_t \rho) / \beta$ are the state coefficients in the feedback function corresponding to the two channels in the feedback function respectively. We refer to $m_t$ as the *weight* of $a_t$ and to $w_t$ as the *weight* of $v_t$; furthermore, let $\pi_t$ be the solution of the following Bernoulli's BVP:

$$\begin{cases} \dot{\pi}_t = \pi_t^2 \left( \vartheta^2 / \kappa + \rho^2 / \beta \right) \ , \\ \pi_T = \gamma \ . \end{cases}$$

We observe that this BVP has the same form as the one obtained for the case of one deterministic channel (44). Moreover, in analogy to the case of Section 3.1, the state coefficients are proportional to the ratio between the marginal productivity parameters of the two channels and their cost (i.e. $\vartheta / \kappa$ and $\rho / \beta$ respectively). Hence, from a deterministic point of view, the optimal policy consists in using more intensely the advertising channel which maximizes this ratio.

**4.2. A risk-associated upper bound.** Let us consider the problem dealt with in the previous Section, assuming that in (39) the cost of $a_t$ becomes negligible (i.e. $\kappa = 0$). Under such hypotheses the problem is ill-posed because the cost matrix $R$ in (39) is singular. This trivial result is the policy called "the-larger-the-better" [**1**], [**7**], [**28**, p.283]: if we assume that in a deterministic maximization problem the activity cost is zero, and the level of activity carried out by the decision-maker increases the utility, then the problem is ill-posed, because there is no trade-off in it.

On the other hand, if we introduce a stochastic channel with a positive risk coefficient, $\sigma > 0$, then the problem remains well-posed even if $\kappa = 0$. In fact, the Riccati equation (40) becomes

$$\begin{cases} \dot{\pi}_t = \vartheta^2/\sigma^2\pi_t + \pi_t^2\rho^2/\beta \ , \\ \pi_T = \gamma \ , \end{cases} \tag{45}$$

and the feedback function (41) is

$$\begin{pmatrix} a_t^* \\ v_t^* \end{pmatrix} = \begin{pmatrix} \vartheta/\sigma^2 \\ \pi_t\rho/\beta \end{pmatrix} \left(\bar{A} - A_t^*\right) \ . \tag{46}$$

Even if the advertising cost of the channel $a_t$ vanishes, there exists an upper bound for the weight function of the control $a_t$. This bound is strictly connected to the risk: the decision-maker has to take into account both the mean evolution and the uncertainty introduced in the system. The policy "the larger, the better" cannot be an optimal choice, because it introduces too much volatility in the state process. Hence, $\vartheta/\sigma^2$ represents the upper bound in the weight function of the control $a_t$ and the concavity of this function is connected with the presence of such a threshold. This observation emphasizes that introducing uncertainty in an advertising channel we deeply modify the structure of the problem. In the following Section we find the optimal solution when one of the two advertising channels is stochastic and we characterize the optimal policy with respect to the risk parameter $\sigma$.

**4.3. Stochastic and deterministic channels.** The most interesting situation occurs when one of the channels affects stochastically the goodwill evolution. Under these assumptions problem (38) has the following optimal feedback control:

$$\begin{pmatrix} a_t^* \\ v_t^* \end{pmatrix} = \begin{pmatrix} \pi_t\vartheta/\left(\kappa + \sigma^2\pi_t\right) \\ \pi_t\rho/\beta \end{pmatrix} \left(\bar{A} - A_t^*\right) = \begin{pmatrix} m_t \\ w_t \end{pmatrix} \left(\bar{A} - A_t^*\right) \ , \tag{47}$$

while the Riccati equation (40) becomes

$$\begin{cases} \dot{\pi}_t = \pi_t^2 \left(\rho^2/\beta + \vartheta^2/\left(\kappa + \sigma^2\pi_t\right)\right) \ , \\ \pi_T = \gamma \ . \end{cases}$$

As $\pi_t$ is a positive and strictly increasing function, it is interesting to study the influence of its values on the feedback evolution.

If the deterministic channel is more efficient than the stochastic one (i.e. $\rho/\beta \geq \vartheta/\kappa$), then it is convenient to invest mainly in the deterministic channel $v_t^*$, because this channel is both more efficient and precise. In fact we have $m_t \leq w_t$ for all $t \in [0, T]$, indipendently of the risk coefficient $\sigma$.

On the other hand, if the stochastic channel is , in mean, more efficient than the deterministic one (i.e. $\rho/\beta < \vartheta/\kappa$), then it may be convenient to use the stochastic channel, but one should consider the risk connected with the control uncertainty. Such a risk can be considered as a cost and it decreases the effectiveness of the control $a_t^*$. In order to stress the dependence of the optimal control on the uncertainty coefficient $\sigma$, let us assume that $\vartheta/\kappa = \lambda\rho/\beta$, with $\lambda > 1$. Furthermore, let $\pi_0 > 0$ be the value at $t = 0$ of the function $\pi_t$ which solves the Riccati equation (we recall that $\pi_0$ depends on the uncertainty coefficient $\sigma$). Comparing the feedback functions $\pi \mapsto \pi\rho/\beta$ and $\pi \mapsto \pi\vartheta/\left(\kappa + \sigma^2\pi\right)$ in the interval $\left[\pi_0, \gamma\right]$, we observe that they take the same value at $\pi^{\#} = 0$ and $\pi^* = (\lambda - 1)\,\kappa/\sigma^2$. Therefore, if $\pi^* \geq \gamma$, the stochastic channel is more intensively used than the deterministic one and this correponds to a small uncertainty coefficient ($\sigma^2 \leq (\lambda - 1)\,\kappa/\gamma$ ). The opposite situation occurs if the uncertainty coefficient is large ($\sigma^2 \geq (\lambda - 1)\,\kappa/\pi_0$), then $\pi^* \leq \pi_0$ and hence the deterministic channel is more intensively used than the stochastic one in the whole programming interval. Finally, for intermediate values of the risk coefficient ($(\lambda - 1)\,\kappa/\pi_0 < \sigma^2 < (\lambda - 1)\,\kappa/\gamma$) it turns out that the stochastic channel, which is more efficient, is more intensively used only in first part of the advertisign campaign, in order to get quickly close to the target goodwill value. On the other hand, in the last part of the programming interval it is useful to control the system with a less effective, but more precise advertising channel, the deterministic one.

We observe that the presence of a stochastic channel makes the situation rather different from the one with deterministic channels only: the two ratios between the marginal productivity and the cost coefficient ($\vartheta/\kappa$ and $\rho/\beta$ respectively) are not sufficient to characterize the optimal policy. The decision-maker must also account for the risk connected with the advertising message.

## 5. Conclusion

In order to understand how two advertising channels can be optimally used in the marketing mix problem for an event planning, we have considered the advertising problem under the assumption that an advertising channel may have a stochastic influence on the goodwill evolution. The objective of the decision-maker is to reach a target goodwill level with the minimum advertising cost. The analysis shows how, in the deterministic framework, the optimal policy depends on the ratio between the productivity and the cost of each advertising channel.

On the other hand, when an advertising channel affects stochastically the goodwill evolution, the decision-maker has to take into account also the risk connected to this control. The model is related to the LQ stochastic optimal control problems which are a subject of increasing interest from both mathematical and economic point of view. The recent mathematical developments ([**1**], [**7**]) in this area represent a very useful toolkit to study some interesting economic problems, that can naturally be described in this framework.

CHAPTER 8

# Results on the marketing model

The main features of this work are connected with two different areas. From an economic point of view, we have dealt with a general model for a new product introduction in the market, formulated using the stochastic control theory. The general problem is hard to solve, and therefore, in order to obtain some explicit results, we introduce the LQ instance of the model. Under this assumption we can use some recent works connected with the theory of LQ control. More in detail, in the literature of the LQ control it is typically assumed that the cost function has a positive definite weighing matrix and the state term has a positive semidefinite weighing matrix. In that case, the solvability of the Riccati equation is both necessary and sufficient for the solvability of the underlying problem. However, stochastic LQ control problems may be well-posed even if the cost matrix and the state weighing matrices are indefinite. This new theory (the recent results on this topic are now appearing in specialized journals) gives some useful instruments which are applied in mathematical finance; with this work we want to suggest that this theory can be successfully applied to other areas of mathematical economics.

We use the stochastic LQ theory to analyze the problem of introducing a new product in the market. Assuming that the decision-maker is risk inclined we can study the two stochastic effects (word-of-mouth publicity and lure/repulsion effect) and we use the closed form solutions obtained in order to explain from an economic point of view the new features of the model.

Moreover, we use the stochastic LQ theory to analyze the communication mix problem for an event planning. Assuming that the decision-maker wants to reach a target goodwill value, we use two different advertising channels to drive the goodwill. In order to compare the efficiency of different advertising channels, we consider the limit case in which one channel affects deterministically the goodwill evolution, while the other does it stochastically. Using the closed form solutions obtained we can explain when it is convenient to use the stochastic channel and when the deterministic one is more efficient.

## 1. Future research developments

The two problems studied in Chapter 6 and 7 are LQ instances of the general problem presented in Chapter 5. We presume that a closed form solution of the general model cannot be obtained. On the other hand, the results given by the solution of the problems in Chapter 6 and 7 surely can be used to obtain some information to transfer to the general problem. However, before considering the general problem under the LQ assumptions, it is worthwhile evaluating the soundness and the applicability of some special problems where we can reach some explicit results. The problems discussed in Chapter 6 and 7 seem to have such characteristics from both mathematical and economic point of view. In fact, we believe that they are the simplest cases and the most natural to consider for this purpose. Further research should be devoted to the study of different special instances of the problem.

An aspect that we have not considered in this work is the possibility of the presence of a budget constraint. The upper bound for the advertising expenditure is usually fixed a priori in planning a publicity campaign. Here, we have a trade-off between the utility and the advertising cost, but the expenditure for the publicity campaign is not upper bounded. A possible direction for further research is the introduction of a new state variable which summarizes all the advertising expenditure. The information carried by this new state variable might be evaluated in a penalty function, which could be part of the objective functional.

Finally, a complete analysis of the Nerlove and Arrow's model, which is made up of two motion equations, one for the goodwill and one for the sales, requires the study of a different aspect. More precisely, it might be interesting to analyze the consequences connected with the introduction of a stochastic goodwill process on the sales evolution. Here, we have not considered such an aspect because it was not relevant for the new product introduction model. However, a general stochastic approach to the marketing problems should account for a stochastic sales process, depending, on one hand, on the stochastic goodwill process and, on the other hand, on different market sources of randomness. Therefore, the analysis of the sales process from a stochastic point of view might be interesting and could be a relevant direction for further research.

# Bibliography


[1] Ait Rami, M., Moore, J.B., Zhou, X.Y., "Indefinite Stochastic Linear Quadratic Control And Generalized Differential Riccati Equation," *SIAM Journal on Control and Optimization*, 40, 4 (published electronically December 7, 2001) 1298-1311.

[2] Buratto, A., "Budget allocation in the Integrated Communication Mix," *Preprint of the Department of Pure and Applied Mathematics of Padua*, 03, February 2002.

[3] Buratto, A., Favaretto, D., "Optimal communication mix to maximize brand image," submitted to *Ricerca Operativa* (December 2001).

[4] Buratto, A., Viscolani, B., "Advertising for the launch of a product with random success," *Ricerca Operativa*, 26 (1996) 45–62.

[5] Buratto, A., Viscolani, B., "New Product Introduction: Goodwill, Time and Advertising Cost," *Mathematical Methods of Operations Research*, 55 (2002) 55-68.

[6] Baldi, P., *Equazioni differenziali stocastiche e applicazioni*, Quaderni dell'Unio ne Matematica Italiana 28, Pitagora Editrice, Bologna, 2000.

[7] Chen, S., Li., X., Zhou, X.Y., "Stochastic Linear Quadratic Regulator With Indefinite Control Weight Costs," *SIAM Journal on Control and Optimization*, 36, 5 (September 1998) 1685-1702.

[8] Feichtinger, G., Hartl, R.F., Sethi, S.P., "Dynamic Optimal Control Models in Advertising: Recent Developments," *Management Science*, 40, 2 (February 1994) 195-226.

[9] Fleming, W.H., Soner, H.M., *Controlled Markov Processes and Viscosity Solutions,* Springer, 1993.

[10] Jørgensen, S., Zaccour, G., "Equilibrium Pricing and Advertising Strategies in a Marketing Channel," *Journal of Optimization Theory and Applications*, 102, 1 (July 1999), 111-125.

[11] Karatzas, I., Shreve, S.E., *Brownian Motion and Stochastic Calculus*, Springer, 1988.

[12] Karatzas, I., Shreve, S.E., *Methods of Mathematical Finance*, Springer, 1998.

[13] Kotler, P., *Marketing Management: Analysis, Planning, Implementation, and Control*, Prentice-Hall, Upper Saddle River, 1997.

[14] Nielsen, L.T., *Pricing and Hedgung of Derivative Securities*, Oxford University Press, New York, 1999.

[15] Nerlove, M., Arrow, J.K., "Optimal Advertising Policy Under Dynamic Conditions," *Economica*, 29 (1962) 129-142.

[16] Raman, K., "Stochastically Optimal Advertising Policies under Dynamic Conditions: The Ratio Rule," *Optimal Control Applications and Methods*, 11 (1990), 283-288.

[17] Rao, R.C., "Extimating Continuous Time Advertising-Sale Models," *Marketing Science*, 5, 2 (Springer 1986), 125-142.

[18] Piccinato, L., *Metodi per le decisioni statistiche*, Springer, Milano, 1996.

[19] Runggaldier, W., "The choice of the loss function for a decision problem under uncertainty as a comporomise between representativity and computational tractability," *International Seminar on Decision Analysis*, Rome (June 18-19 2001), 141-148.

[20] Sethi, S.P., "Deterministic and Stochastic Optimisation of a Dynamic Advertising Model," *Optimal Control Applications and Methods*, 4, 2 (June 1990), 179-184.

[21] Tapiero, C.S., "On-Line and Adaptive Optimum Advertising Control by a Diffusion Approximation," *Operations Research*, 23, 5 (September-October 1975), 890-907.

[22] Tapiero, C.S., "Optimum Advertising and Goodwill under Uncertainty," *Operations Research*, 26, 3 (May-June 1978), 450-463.

[23] Tapiero, C.S., "A Generalization of the Nerlove-Arrow Model to Multi-Firms Advertising under Uncertainty," *Management Science*, 25, 9 (September 1979), 907-915.

[24] Tapiero, C.S., "Optimal Control of a Stochastic Model of Advertising," in Feichtinger, G. (ed.), *Optimal Control Theory and Economic Analysis*, North-Holland, Amsterdam, 1982, 287-299.

[25] Tapiero, C.S., "Stochastic Diffusion Model with Advertising and Word-of-Mouth Effects," *European Journal of Operations Research*, 12 (1983), 348-356.

[26] Tapiero, C.S., *Applied Stochastic Models and Control in Management*, North-Holland, Amsterdam, 1988.

[27] Varian, H.R., *Microeconomic Analysis*, Norton, New York, 1984.

[28] Yong, J., Zhou, X.Y., *Stochastic Controls, Hamiltonian Systems and HJB Equations*, Springer, New York, 1999.