FINANCIAL MODELING

A Backward Stochastic Differential Equations Perspective

S. Crépey, Springer Finance Textbooks, June 2013

Abstracts by Chapters

1 Some Classes of Discrete-Time Stochastic Processes

The most important mathematical tools in pricing and hedging applications are certainly martingale and Markov properties. Martingality can be stated as a stochastic equation written in terms of conditional expectations. The Markov property can be stated in terms of deterministic semi-group Kolmogorov equations. The tower rule for conditional expectations and Markovian semi-group equations can be considered as primary dynamic programming pricing equations, in their stochastic and deterministic form (to mature as stochastic pricing BSDEs and deterministic pricing PIDEs later in the book). Practically speaking, a Markov property is a necessary companion to a martingale condition in order to ensure tractability (up to "the curse of dimensionality"). From an opposite perspective, martingales can be used to pursue some theoretical developments beyond a Markov setup. It's interesting to note that historically, stochastic calculus was first developed as a tool for the study of Markov processes, until it was realized that the theory of martingales, in particular, was interesting for its own sake, allowing one to supersede Markovianity.

To emphasize these fundamental ideas at the simplest possible level of technicality, we present them in this first chapter in discrete time. We start with the definition and main properties of conditional expectation. The former is in fact mainly useful to prove the latter (which we leave to standard probability textbooks), since most practical conditional expectation computations are directly based on these properties, with the tower rule as a most emblematic example. We then introduce discrete time Markov chains and martingales in this spirit and with finance already in mind – basing most of our examples on random walks fortune processes, like with the doubling strategy, which grants a wealth of one for sure in the end but with values that are unbounded from below on the way!

2 Some Classes of Continuous-Time Stochastic Processes

So far we have studied random processes in discrete time. We now turn to studying random processes in continuous time, for which the unformal definitions of filtration, conditional expectations, martingales, submartingales, supermartingales, stopping times, Markov processes... are essentially the same as in discrete time – but continuous-time entails some technicalities! By the way did you first believe, judging by the names, that a submartingale should be nondecreasing on average and a supermartingale nonincreasing? If not, that's because you forgot to turn your head and look backward, i.e. we want that a submartingale and a supermartingale attached to the same terminal condition (random variable ξ) are "in the right order" (so sub- under super-, as should be). Did you forget that we have financial derivatives in mind, which are defined in terms of a terminal payoff ξ at a future time point (maturity) T and will be studied later in the book through backward SDEs?

3 Elements of Stochastic Analysis

Our purpose in this chapter is to give an overview of the basics of stochastic calculus, an important mathematical tool that is used in control engineering, in modern finance and insurance, and in modern management science, among other fields. The chain rule of stochastic calculus, the so called Itô formula, is one of the most used mathematical (or probabilistic) formulas in the world, since it implicitly sits under every trader's screen. At least the Itô formula gives rise to one of the mostly posed mathematical exercises, as follows. Let $dS_t = S_t \sigma dW_t$, starting from today's observed value for S_0, model the returns of a stock price, where W_t is a Brownian motian and σ is the so called volatility parameter (the "temperature" on financial markets). What are the dynamics for $X_t = \ln(S_t)$? But this is the celebrated Black-Scholes model! Then we will add jumps, to make it more spicy and because in this book the Brownian motion W_t and the Poisson process N_t are equally treated on a fair basis, as the prototype and the fundamental driver of continuous and jump processes, respectively. Now, as opposed to the above forward SDE, endowed with an initial condition for S_0 at time 0, it's now time to consider our first backward SDE. That's because derivative contracts are defined in terms of a payoff ξ at a future maturity T. This payoff ξ is random and defined in terms of an underlying such as S_T, but what we are looking for is the price and the hedge of the derivative at the current pricing time $t < T$. These price Π_t and hedge Δ_t are obtained as the solution of a backward SDE such as $d\Pi_t = \Delta_t dS_t, \Pi_T = \xi$. The solution of a BSDE has therefore two components, Π and Δ. In case of American options with early exercise clauses, there is a third component A, intended to maintain the value process Π above the payoff. Otherwise isn't a BSDE too simple and directly solved by the application of a suitable martingale representation theorem? But that's only because we forgot the issue of funding our position. Funding costs give rise to an additional term $g_t(\Pi_t, \Delta_t)dt$ in the BSDE. Moreover, since the crisis, in nowadays market environments, funding costs involve some nonlinearities (such as two different rates for lending and borrowing, if you are a risky bank). Dealing with such nonlinearities is precisely what BSDEs were invented for.

4 Martingale Modeling

In this chapter, we show how the task of pricing and hedging financial derivatives can be reduced to that of solving related backward stochastic differential equations, called stochastic pricing equations in this book, equivalent to the deterministic pricing equations that arise in Markovian setups. The deterministic pricing equations, starting with the celebrated Black-Scholes equation, are better known to practitioners. However, these deterministic partial-differential equations, also including integral terms in models with jumps, are more "model dependent" than the stochastic pricing equations. Moreover, the deterministic pricing equations are less general since they are only available in Markovian setups. In addition, the mathematics of pricing and hedging financial derivatives is simpler in terms of the stochastic pricing equations. Indeed, rigorous demonstrations based on the deterministic pricing equations involve technical notions of viscosity or Sobolev solutions (at least as soon as the problem is nonlinear, e.g. when early exercise clauses and related obstacles in the deterministic equations come into the picture).

5 Benchmark Models

In this chapter we give a very succint primer of basic models for reference derivative markets: equity derivatives (Black-Scholes model and stochastic volatility or/and jump extensions), interest rate derivatives (Libor market model) and credit derivatives (one-factor Gaussian copula model). Can we really call these models? They are more quotation devices, used by traders for conveying the information regarding value of derivatives in units of measurements with a more physical significance than simply a number of dollars or euros. We want to speak of the implied volatility, which gives the "temperature" of the markets at a given level of maturity and strike for an option, and of implied correlation, which is the measure of (credit) "contagion".

More serious (calibratable, see Chapter 9) models are affine models (with affine drift, covariance and jump intensity coefficients). In these models, semi-explicit Fourier formulas are available for the prices and Greeks of European vanilla options. Yes, there is a lot of jargon in this theory, such as "Greeks" in reference to the letters (e.g. Δ above) which are used for various risk sensitivities, and European (or American above, but also Russian, Parisian, and somewhere in the book I contrived "Hawaiian"), or vanilla (as opposed to exotic), in reference to different kinds of options (remember, this is a global world). In Chapter 9 we come to understand why we need such fast pricing formulas at the stage of model calibration, which involves intensive pricing of vanilla options. However, as far as pricing exotics or dealing with less standard models is concerned, the pricing equations have to be solved numerically, which is the object of the next few chapters.

6 Monte Carlo Methods

The term "Monte Carlo" for computational methods involving simulated random numbers was invented by Metropolis and Ulam when they were working at the Las Alamos Laboratory (didn't some people say that quants were rocket scientists?). Like deterministic pricing schemes, simulation pricing schemes can be used in any Markovian (or, of course, static one-period) setup. In the case of European claims, simulation pricing schemes reduce to the well known Monte Carlo loops. For products with early exercise features, or for more general control problems and the related BSDEs, numerical schemes by simulation are available too, yet these are more sophisticated and will be dealt with separately in Chapters 10 and 11.

Monte Carlo methods are attractive by their genericity: genericity of their theoretical properties (such as the confidence interval they provide for the solution, at least for genuine pseudo Monte Carlo methods, as opposed to the quasi Monte Carlo methods also reviewed in this chapter); and also genericity of implementation. But (pseudo) Monte Carlo methods are slow, only converging at the rate $\sigma m^{-\frac{1}{2}}$, where m is the number of simulation runs and σ is the standard deviation of the sampled payoff. So ultimately with Monte Carlo it's all about how to make it faster. To accelerate the convergence, various variance reduction techniques (e.g. control variate and importance sampling) can be used to transform a given payoff into another one with less variance. An alternative to variance reduction is quasi Monte Carlo, which converges faster in practice than pseudo Monte Carlo (but beware of the dimension; moreover quasi Monte Carlo estimates do not come with confidence intervals). A last "acceleration" technique is of course to resort to a parallel implementation, which with Monte Carlo is an easy thing to do, but unfortunately parallelization techniques are not dealt with in this book!

7 Tree Methods

Tree pricing schemes are natural in finance because of their Markov chain interpretation as discrete time pricing models. From a practical point of view, trees are often rather obsolete as compared with more sophisticated finite difference or finite element technologies. However, in a number of situations, they remain an adequate and simple alternative. Moreover, from the theoretical point of view, the Markov chain interpretation underlies interesting probabilistic convergence proofs of the related (deterministic) pricing schemes.

Note that there is no hermetic frontier between deterministic and stochastic pricing schemes. In essence, all these numerical schemes are based on the idea of propagating the solution, starting from a surface of the time-space domain on which it is known (the maturity of the derivative), along suitable (random) "characteristics" of the problem (here "characteristics" refers to Riemann's method for solving hyperbolic first-order equations). From the point of view of control theory, all these numerical schemes can be viewed as variants of Bellman's dynamic programming principle. Monte Carlo pricing schemes may thus be regarded as one-time-step multinomial trees, converging to a limiting jump diffusion when the number of space discretization points (tree branches) goes to infinity. The difference between a tree method in the usual sense and a Monte Carlo method is that a Monte Carlo computation mesh is stochastically generated and nonrecombining.

8 Finite Differences

This book emphasizes the use of backward stochastic differential equations (BSDEs) for financial modeling. The BSDE perspective is useful for various reasons. This said, in Markov setups, BSDEs are nothing than the stochastic counterparts of the more familiar, deterministic, partial differential equations (PDEs, or PIDEs for partial integro-differential equations if there are jumps), for representing prices and Greeks of financial derivatives (have you already heard of the Blach-Scholes equation?). This means that prices and Greeks can also be computed by good old deterministic numerical schemes for the corresponding P(I)DEs, such as finite differences (for pricing problems which are typically posed on rectangular domains, the additional complexity of using potentially more powerful finite element methods is often not justified). Like tree methods and as opposed to simulation methods, finite difference methods can easily cope with early exercise features, and in low dimension they can give an accurate and robust computation of an option price and Greeks (delta, gamma, and theta at time 0). However, they are not practical for dimension greater than three or four, for then too many grid points are required for achieving satisfactory accuracy.

9 Calibration Methods

Prices of liquid financial instruments are given by the market and are determined by supply-and-demand. Calibrating a model means finding numerical values of its parameters such that the prices of market instruments computed within the model, at a given time, coincide with their market prices. Liquid market prices are thus actually used by models in the "reverse-engineering" mode that consists in calibrating a model to market prices. Once calibrated to the market, a model can be used for Greeking and/or for pricing more exotic claims (Greeking means computing risk sensitivities in order to set-up a related hedge).

Calibration thus corresponds to estimation of a model. However, in finance the term "estimation" specifically refers to statistical estimation, i.e. estimation based on historical data by maximum likelihood or any other statistical procedure. Statistical estimation is thus backward looking, whereas calibration is forward looking, since derivative prices at the current time are based on the views of the market regarding the future dynamics of the underlyings. It is generally acknowledged that, whenever option data are available, it is better to use them to calibrate the model than to estimate a model statistically on past data.

The simplest example of a calibration problem is encountered in Chapter 5, where we discuss the notions of the implied volatility of an option and the implied correlation of a CDO tranche. In these cases the calibration problem is easy since there is only one parameter to "calibrate" to only one market quote. But can this really be called calibration? Well, it depends on what one wants to do. The exercise for the bank is, given a product it's interested in, to identify a number of risk factors, select a number of hedging instruments also responsive to these, and to devise a model consistent for everybody (derivative and its hedging assets, which in some cases can be derivatives themselves) such that the risk of the position can be monitored in this model. But wait! where has the real world gone? the model only exists in our heads until is is calibrated to the market. The least requirement is that the current price of the derivative and its hedging assets in the model are consistent with the ones oberved today in the market. For instance, if it is only about hedging a vanilla option with the underlying stock, calibration in the sense of fitting a Black–Scholes implied volatility to that option's price may be enough. If there are other derivatives among the hedging assets, then it's not only a matter of calibration to one price, but of co-calibration of the whole set of instruments in use. Moreover (co-)calibration at a given time is only a first step, which can always be achieved in simple models such as local volatility models. These tell us that the volatility of a stock S is random (which it is), but only as a function of S. This would mean that options can be perfectly hedged by their underlying. Do you really believe in this? You shouldn't, since it would simply contradict the existence of derivative markets (which would be useless if derivatives could be synthetized in terms of their underlyings). What we are missing here is the dynamics, namely we need a (co-)calibratable model with, in addition, the right dynamics, or the right Greeks. Now there is a "meta-theorem" in financial modeling stating that right Greeks are stable Greeks, namely Greeks which are stable when the model is recalibrated to the market every day. Which is a matter of calibration again, but across time (stability of the recalibrated parameters). So: calibration, co-calibration and re-calibration, the master equation of financial modeling!

10 Simulation/Regression Pricing Schemes in Diffusive Setups

We reviewed simulation and deterministic pricing schemes in Chapters 6 through 8. Analogies and differences between simulation and deterministic pricing schemes are most clearly visible in the context of pricing by simulation claims with early exercise features (American and/or cancelable claims). Early exercisable claims can be priced by hybrid "nonlinear Monte Carlo" pricing schemes in which dynamic programming equations, similar to those used in deterministic schemes, are implemented on stochastically generated meshes. Such hybrid schemes are the topics of Chapters 10 and 11, in diffusion and in pure jump setups, respectively.

In Chapter 10 we deal with the issue of pricing convertible bonds numerically, by simulation. A convertible bond can be regarded as a coupon-paying and callable American option. Moreover, call times are subject to constraints, known as call protections, preventing the issuer from calling the bond at certain sub-periods of time. The nature of the call protection may be very path-dependent, leading to high-dimensional nonlinear pricing problems. Deterministic pricing schemes are then ruled out by the curse of dimensionality, and simulation methods are the only viable alternative. The numerical results of this chapter illustrate the good performances of the simulation regression scheme for pricing convertible bonds with highly path-dependent call protection. More generally, this chapter is an illustration of the power of simulation approaches, which automatically select and loop over the most likely future states of a potentially high-dimensional, but also often very degenerate, factor process X, given an initial condition $X_0 = x$. By contrast, deterministic schemes loop over all the possible states of X_i regardless of their likelihood. Simulation schemes thus compute "where there is light". By contrast, deterministic schemes compute everywhere, also in "obscure", useless regions of the state space. In the context of path-dependent payoffs with a high-dimensional but very degenerate factor process, a simulation scheme thus exploits the degeneracy of a factor process, whereas this path-dependence makes the deterministic scheme impractical.

11 Simulation/Regression Pricing Schemes in Pure Jump Setups

In this chapter we devise simulation/regression numerical schemes in pure jump models. There the idea is to perform the nonlinear regressions, used for computing conditional expectations, in the time variable for a given state of the model rather than in the space variables at a given time in the diffusive setups of Chapter 10. This idea is stated in the form of a generic lemma that is valid in any continuous-time Markov chain model. This is then tested in the context of two credit risk applications, the first of which values the sensitivities of a CDO tranche in a homogeneous groups model of portfolio credit risk by Monte Carlo without resimulation. The second computes by Monte Carlo the CVA on a CDO tranche in a common shock model of counterparty credit risk. CVA stands for credit valuation adjustment, the correction in value to a derivative accounting for the default risk of your counterparty, a topical issue since the crisis. But wait: are you perfect yourself? Isn't it so that most Western banks nowadays quote at a few hundreds of basis points of credit spread? This means that you should also account for your own default risk in the valuation, otherwise I doubt many clients would agree to deal with you – which implies the related nonlinear funding struggle that if you are credit risky, the funding of your position will involve (at least) two rates, a lending and a borrowing one. Now, quiz to the reader (not answered in this chapter, and in fact nowhere else either): how would you price nonlinear funding costs on a very high-dimensional and discrete underlying like a CDO tranche?

12 Backward Stochastic Differential Equations

We saw in Chapter 4 that the problem of pricing and hedging financial derivatives can be modeled in terms of (possibly reflected) backward stochastic differential equations (BSDEs) or, equivalently in the Markovian setup, by partial integro-differential equations or variational inequalities (PIDEs or PDEs for short). Also, chapters 10 and 11 just provided thorough illustrations of the abilities of simulation/regression numerical schemes for solving high-dimensional pricing equations: very large systems of partial differential equations in Chapter 10 and Markov chain related systems of ODEs in Chapter 11.

Now that we experimented the power of the theory, let's dig into it. The next few chapters provides a thorough mathematical treatment of the BSDEs and PDEs that are of fundamental importance for our approach. More precisely, Chapters 12 to 14 develop, within a rigorous mathematical framework, the connection between backward stochastic differential equations and partial differential equations. This is done in a jump-diffusion setting with regime switching, which covers all the models considered in the book. To start with, Chapter 12 establishes the well-posedness of a Markovian reflected BSDE in a rather generic jump-diffusion model with regime switching, denoted by (X, N), which covers all the models considered in this book. In standard applications, the main component of the model, in which the payoffs of a derivative are expressed, is X. The other model component N can be used to represent a pricing regime, which may also be viewed as a degenerate form of stochastic volatility. More standard diffusive forms of stochastic volatility may also be accounted for in X. The presence of jumps in X is motivated by the empirical evidence of the short-term volatility smile in the market. In credit and counterparty risk modeling, the main model component (the one which drives the cash flows) is the Markov-chain-like-component N, representing a vector of default status and/or credit ratings of reference obligors; a jump-diffusion-like-component X can be used to represent the evolution of economic variables modulating the dynamics of N. Frailty and default contagion are accounted for by the coupled interaction between N and X.

13 Analytic Approach

In this chapter we derive the companion variational inequality approach to the reflected BSDEs of Chapter 12. First we introduce systems of partial integro-differential variational inequalities associated with these BSDEs and we state suitable definitions of viscosity solutions for related problems. Remember that BSDEs are used to model nonlinear phenomena, meaning that the equivalent PDEs (or systems of them, or PIDEs) are nonlinear too. They therefore don't have classical solutions, but only solutions in weaker senses, viscosity solutions being the notion of choice for the kind of nonlinearities we face in pricing (or more general control) problems, which at least have some kind of comparison property (recall the sub- versus super-martingale story that we sketched in the discussion of Chapter 2).

We then deal with the corresponding existence, uniqueness and stability issues. The value processes (first components) in the solutions of the BSDEs is characterized in terms of the value functions for related optimal stopping or Dynkin game problems. We then establish a discontinuous viscosity solutions comparison (again) principle, which is the deterministic counterpart of the BSDEs comparison theorem alluded to above. In particular, this comparison principle implies uniqueness of viscosity solutions for the related obstacle problems. The comparison principle is also used for proving the convergence of stable, monotone and consistent deterministic approximation schemes. The notion of viscosity solutions is nice because everything happens as if it wasn't there: all the classical results which apply to linear problems can be extended to nonlinear problems endowed with a comparison property, provided one switches to the notion of viscosity solutions for these problems. But the underlying mathematics are nontrivial, which is why we need Chapter 13!.

14 Extensions

Have you ever seen a financial contract which only pays one single cash-flow at a terminal maturity T? There are some (I heard "vanillas"!), but not so many. Ever seen a continuous-time coupon stream? I'd be surprised. More broadly, to account for many real-life practical features such as discrete dividends or discrete path-dependence, we need to extend the theory a little bit. With this motivation, in this chapter, we provide various extensions to the BSDE and PDE results of Chapters 12 and 13. First we deal with discrete dividends on a financial derivative or an underlying asset. Then we extend the results of the previous chapters to more general reflected BSDEs that appear in the case of intermittent call protection in Chapter 10.

15 Technical Proofs ($**$)

In this chapter we provide the proofs of the most demanding results of Chapters 12 through 14. Have you noticed the ($**$) in the title? It means "very difficult". Given the dual nature of the proposed audience for the book (scholars and quants), I have provided in the first chapters a lot of background material. Yet I didn't want to avoid the sometimes difficult mathematical technique that is needed for deep understanding. So, for the convenience of readers, we signal sections that contain advanced material with an asterisk (*) or even a double asterisk (**) for the still more difficult portions.

16 Exercises

It is well known that one learns through practice, so exercises and problems, but it's also good if one can chek one what did. So, exercises and problems: corrected or not? Let's go for both. This chapter is devoted to exercises related to the probabilistic and stochastic analysis background of Chapters 1 through 3. They are not corrected (solutions of the exercises are only available for course instructors). However I must say most of them are very classical exercises.

17 Corrected Problem Sets

Chapter 17 provides problem sets for the financial and numerical material of Chapters 4 through 9. The solutions as well as a few Matlab scripts are downloadable from the website of the book. The Matlab scripts are truly worth being run if you want to really understand what happens in the numerics.