1.0 Scenario Generation

The previous sections described that, given a set of scenarios which fully describe a trading universe, numerical methods from stochastic and linear/nonlinear programming can be used to formally obtain an optimal replication or hedge portfolio for a given benchmark. This section takes a closer look at the scenario generation process, which is the most important support for the optimization framework. The subject is very broad, and this discussion is restricted to provide a “road map” of implementation difficulties with particular emphasis placed on numerical issues. The spectrum of the approach is fully identified however.

1.1 Formal problem

In previous discussions, we assumed the existence of a vector of present values (PV) describing a portfolio, given by

\[ \hat{q} = [q_1 \ q_2 \ \ldots \ q_T]^T. \]  

(1.1)

The components of this vector can represent PV of arbitrary securities. Computing a single element in the vector is done by means of arbitrage-free models discussed earlier on in the course. These include Black-Scholes-based methods for pricing equity derivatives, and the Heath-Jarrow-Morton (HJM) framework for pricing instruments with a dependence on interest rates. Numerical methods were presented to solve these problems, among which finite difference methods and trees (lattices). Scenario generation is essentially a tree approach. Also, HJM provides at present one of the most general framework for pricing securities, so it will be used here for illustration (Jarrow 1996).

Figure 1-2 gives the “location” of the current discussion in terms of the economy approximation diagram of the HJM framework, discussed in previous lectures. It is important to note that the map between a continuous pseudo economy and a discrete one requires a choice of a parametric representation for the forward rate, and depending on this, an approximation for the probability distribution function associated with the martingale measure. The restriction on these parametric choices is that they should converge to the true forward rate expression as the time steps go to zero.

In the pseudo discrete framework, a tree is used to represent the stochastic evolution of interest rates and of other market factors. The tree is initially built on portfolios of zero-coupon bonds (“feasibility mode”), with as many bonds as there are time steps in the
discretization. The tree is then used to replicate cash flows of derivative securities with the
bonds as underlyings. This involves propagating the price vector backward and calculate
an expected value at every step. This process starts at the terminal nodes, at which the value
of the option is known, and proceeds by backward induction to result in the PV of the option
at the present time. For certain types of path dependent securities (such as Asian options),
there is no closed form expression for the probability distribution function used in this step.,
so pricing this type of instruments requires a parametric approximation here as well.

A representation of this tree is given in Figure 1-3, illustrating two more parameters that

Figure 1-3. Representation of a tree for pricing securities

enter the formulation: the representation of the state space as a finite number \( m \) of market
factors, yielding \( m + 1 \) branches at every node in the tree, and the discretization of the time

interval into \( n \) time steps. The finer the time discretization and the representation of the state space, the better the approximation of the discrete formulation.

1.2 Correlation of market factors

In a general case, the instruments entering a portfolio used for hedging/replication depend on market factors which are mutually correlated. The following is an overview on how correlations can be handled in a general way. The discussion is taken from (Amin and James N. Bodurtha 1995), in which the problem of valuing a path-dependent American option depending on two term structures and one exchange rate is described. The approach can be easily generalized to include a larger number of market factors and/or different types of securities (options on equities, etc.).

The approach consists of defining stochastic processes for a domestic interest rate (market factor number one), a foreign interest rate (market factor number two), and the exchange rate between foreign and domestic currencies (market factor number three), all of which are correlated. The discrete parametric forms for each of these processes are chosen so to ensure convergence to the forward rates established by the continuous time HJM model, in which the conditions of arbitrage-free and completeness of the market are assumed. This guarantees the existence of a unique martingale measure \( Q \), which in turn allows any contingent claim to be written as an expectation of its payoff discounted by a spot rate of interest.

The volatility function which enters the formulation is arbitrary, and the discrete representation of the model is a full tree with four branches at every node. If the volatility is assumed to be constant, or if interest rates are assumed to be normally distributed, the valuation tree has an isomorphic lattice representation (in the first case, the Ho-Lee model is recovered).

The time to maturity is assumed to be discretized as shown in Figure 1-4, in which \( \Delta \) represents a unit time interval, \( \tau \) is the maturity of the option, and \( T \) is the maturity of the longest-lived underlying.

The domestic interest rate is described by the stochastic difference equation

\[
 f_d(t + \Delta, T) - f_d(t, T) = \alpha_d(t, T; \tau)\Delta + \sigma_d(t, T; \tau)X_d(t + \Delta)\sqrt{\Delta},
\]

where the terms in the right-hand side represent the drift and variance of the forward rate. \( s_i \) are the node indices at each layer in the tree. \( \{X_d(i\Delta)\}_{i=1,2,...} \) represents a sequence of independent normal random variables under the martingale probability measure \( Q \). The spot interest rate \( r_d(t) \) is defined as \( r_d(t) = f(t, t) \).
The foreign interest rate is described in exactly symmetric fashion,

\[ f_j(t + \Delta, T) - f_j(t, T) = \alpha_f(t, T; s_t) \Delta + \sigma_f(t, T; s_t) X_j(t + \Delta) \sqrt{\Delta}, \]  

(1.6)

with the correlation between \( X_d(t) \) and \( X_j(t) \) given by \( \rho_{df}(t, s_t) \). The spot exchange rate in domestic currency units is described by

\[ \ln \frac{S(t + \Delta)}{S(t)} = (\alpha_s(t, s_t) + r_d(t) - r_f(t)) \Delta + \sigma_s(t, s_t) X_s(t + \Delta) \sqrt{\Delta}, \]  

(1.7)

and the covariance matrix of the random variables is given by

\[ \text{cov}(X_d(t), X_f(t), X_s(t)) = \begin{bmatrix} 1 & \rho_{df}(t, s_t) & \rho_{ds}(t, s_t) \\ \rho_{df}(t, s_t) & 1 & \rho_{fs}(t, s_t) \\ \rho_{ds}(t, s_t) & \rho_{fs}(t, s_t) & 1 \end{bmatrix}. \]  

(1.8)

The method proceeds next by describing a set of four elementary securities, shown in , with

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig1-9.png}
\caption{Elementary trading securities}
\end{figure}

which all other securities in the economy can be described using a finite number of trades. The discounted prices\(^1\) are then obtained, and a risk neutral valuation argument is applied to them to eliminate the drift terms from the forward rate expression. In functional form, the drift terms are expressed as

\[ \alpha_d = \alpha_d(\sigma_d, X_d, t, T) \]
\[ \alpha_f = \alpha_f(\sigma_f, \sigma_s, X_s, \alpha_s, t, T) \]
\[ \alpha_s = \alpha_s(\sigma_s, X_s, t, T) \]

(1.10)

and are thus solely dependent on the volatility and random terms. The forward rates can therefore be expressed as

\[ f_d = f_d(\sigma_d, \rho_{df}, \rho_{ds}, X_d, t, T, s_t) \]
\[ f_f = f_f(\sigma_s, \sigma_f, \rho_{df}, \rho_{fs}, X_s, \alpha_s, t, T, s_t) \]
\[ S(t) = S(\sigma_s, \rho_{ds}, \rho_{fs}, r_d, r_f, X_s, t, T, s_t) \]

(1.11)

\( \text{ } \)

---

1. The discounted price of a security is defined as \( Z(t, T) = \frac{P(t, T)}{B(t)} \), where \( P(t, T) = \prod_{j=i}^{T-1} f(t, j) \)
and depend only the volatility and correlation inputs. The asymmetric representation of the foreign forward rate with respect to the other two processes is due to the conversion of all expressions into a reference domestic currency.

1.3 Scenario generation

The above discussion can be generalized to a multiple market factor framework. Assuming that the original portfolio of the first section,

\[ \hat{q} = \begin{bmatrix} q_1 & q_2 & \cdots & q_I \end{bmatrix}^T \]

contains instruments with price depending on \( m \leq I \) market factors. The portfolio can be decomposed into \( m \) subsets, as shown in Figure 1-12. Going back to the tree representation

\[ \hat{q} = \begin{bmatrix} q_1, \cdots, q_{k1}, q_{k1+1}, \cdots, q_{k2}, \cdots, q_{ki+1}, \cdots, q_{k(m-1)}, \cdots, q_{km} \end{bmatrix}^T \]

depends at least on factor:

\[ \sum_{i=1}^{m} k_i = I \]

Figure 1-12. Portfolio partitioning as a function of market factors

given in Figure 1-3, a scenario is then simply a path in the tree, from root node to some terminal node. This is summarized in Figure 1-13. The probability of each scenario is given

\[ p_{s_j} = \begin{bmatrix} p_1 \\ \vdots \\ p_{m+1} \end{bmatrix} \]

\[ (m+1)^{n+1} - 1 \] nodes in the tree

\[ (m+1)^n \] terminal nodes

Figure 1-13. General scenario tree

by the product of the probability of every branch along the path,
For $n$ time steps and $m$ market factors, there are $(m + 1)^n$ distinct paths, or scenarios, in the tree.

1.4 Numerical implementation issues
The preceding section described a scenario generation methodology in the HJM framework, and this represents by far the simplest approach: other models either do not yield a stochastic evolution of term structures, so each process needs to be calibrated at every time step, or assume a dependence on the drift terms which increases the parameter space.

Implementation difficulties of the above can be grouped into two distinct classes, summarized as follows:

A. Modeling (conceptual) issues

- Discretization of the forward rate. A challenge is to find a parametric form which maximizes the rate of convergence of the model.

- Approximation for the pdf of the martingale measure $\tilde{Q}$, if necessary (for example, there is no closed form expression for the pdf of an arithmetic Asian option, and six approximations are currently used in practice).

- Optimal representation of the state space. This consists in selecting a significant number of market factor while avoiding redundancy. In the HJM framework, the state space is fully described by the forward rate stochastic process.

B. Computational (feasibility) issues

- Discretization of the time horizon. The accuracy of scenario generation increases with a finer discretization, but the number of scenario increase exponentially with the number of time steps. Since most of the decisions in the tree (i.e., exercising an option) occur at specific dates, it is possible in practice to make wise choices for the discretization, for example by increasing the sampling near predefined cash flow dates (coupon payments of a bond, etc.).

- Description of the state space as a minimum number of market factors. The number of branches at every node likewise increases the number of scenarios. Approximations used in practice include the relaxation of stochastic volatility requirements. By assuming a deterministic form for the volatility, the tree can be mapped into a lattice (and the number of paths is now linear in the number of market factors and time steps).

- Generation of the $D$ matrix of future prices. This step can be compute-time intensive, depending on the type of instruments being priced.
• Estimation of the covariance matrix of term structures. The volatility is not a directly observable quantity, and its estimation is not well-understood conceptually. One may rely on historical estimates, as described in (JPMorgan/Reuters 1996), or on an estimate implied from a pricing model. Both methods have their limitations. See section 6.3 of (Musiela and Rutkowski 1997) for a general discussion.

Given the above difficulties, it is not surprising that scenario generation is often done in an ad-hoc fashion in practice, with the selection of both scenarios and probabilities attached to it representing the simplest (and most dangerous) alternative to the formal approach described here. Finally, it should be noted that scenarios generation methods based on a single period can be very sensitive to forecasting errors, and also represent a dangerous choice.

Improvements based on simple “scenario picking” methods have been proposed. The following “rules of thumb” are often applied in practice (Dembo 1994):

• Choose a uniform distribution for the probability of all underlying instruments.
• Generate scenarios corresponding to up and down shifts in the volatility smile (dependence of the volatility of an option price on its strike price).
• Generate (parallel or proportional) shifts in the term structure of interest rates.
• Add a few “catastrophic” scenarios with low probabilities to represent jumps in variables.

2.0 Monte Carlo simulation

Simulation can be used to generate scenarios. What follows is a brief reminder of Monte Carlo simulation methods.

2.1 General

Monte Carlo simulation consists in making use of random number generators to produce sample points of a function which is either analytically intractable or which is too cumbersome to evaluate by other numerical methods. The formal problem is posed as follows. See Chapter 8 of (Lamberton and Lapeyre 1996).

For a random variable with a distribution law given by , the goal is to generate a sequence of independent trials , all with the same distribution , and relying on the law of large numbers

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(X_n) = \int f(x) \mu(x) dx
\]  (2.1)
for convergence ($f(x)$ is some $\mu$-integrable function). $N$ is the sampling parameter. There are three distinct steps in applying a Monte Carlo simulation algorithm:

1. **Seed generation:**

Generate a sequence of pseudo-random numbers

$$\{U_n\}_{n \geq 1},$$

the realization of a sequence of independent uniformly distributed random variables on the closed interval $[0, 1]$.

2. **Choice of function:**

Find a function $F$ so that $F(\{U_p\})$ follows the distribution law $\mu dx$ for some natural number $p$.

3. **Construction of the new random variable:**

Generate the sequence $\{X_n\}_{n \geq 1}$ with $X_n = F(\{U_k\}_{k \leq p})$, where $\{U_k\}_{k \leq p}$ is some (not necessarily strict) subsequence of $\{U_p\}$.

The following example provides a simulation of a normally distributed random variable with mean $m$ and variance $\sigma$:

1. Generate $\{U_n\}_{n \geq 1}$ using some pseudo-random number generator. This is a sequence of independent, uniformly distributed random variables on $[0, 1]$.

2. Find the function $F$. It can be shown that

$$f(U_1, U_2) = \sqrt{-2\log(U_1)\sin(2\pi U_2)}$$

is a random variable which is normally distributed with mean zero and variance one (henceforth a standard Gaussian variable). This choice is not unique. Hence define $F(U_1, U_2) = m + \sigma f(U_1, U_2)$.

3. Construct the sequence $\{X_n\}_{n \geq 1}$ with $X = F(U_1, U_2)$.

This can be extended to the simulation of a Gaussian vector with mean $\hat{m} = E((X_1), ..., E(X_d))$ and with covariance matrix $G = [\sigma_{ij}]_{d \times d}$, where

$$\sigma_{ij} = E(X_iX_j) - E(X_i)E(X_j).$$

By construction, $G$ is symmetric positive definite, but it is also assumed to be invertible (this assumption is often not satisfied in practice). A pre-
treatment step also consists in finding the Cholesky decomposition of $G$, i.e., finding the lower triangular matrix $A$ such that $AA^T = G$. An algorithm to find the Cholesky factorization is given in Section 2.3. This decomposition method is not unique, and the same “square root” matrix could be obtained from principal component/eigenvector analysis, but using a lower triangular matrix reduces the number of computations (note: if $G$ is invertible, so is $A$).

1. Construct by simulation $d$ standard Gaussian variables
   $$\hat{Q} = (q_1, \ldots, q_d),$$ as outlined above.

2. Observe that $\hat{Z} = A^{-1}(\hat{X} - \hat{m})$ is a Gaussian vector of mean zero and variance $I$ (from now on, a Gaussian random vector).

3. Compute $\hat{X} = \hat{m} + A\hat{Q}$.

Monte Carlo methods can be used to simulate a stochastic process, such as a Wiener process (or geometric Brownian motion) $\{W_t\}_{t \geq 0}$. One way to achieve this is to renormalize a sequence $\{X_i\}_{i \geq 0}$ of random walks which are independent identically distributed (iid), and with distribution law given by

$$p(X_i) = \begin{cases} 1 & \text{if } X_i = 1 \\ -1 & \text{if } X_i = -1 \end{cases} \quad (2.3)$$

This ensures that the random walks are standard Gaussian variables, which can be generated as described earlier. The function which takes the simulated random walks as argument is defined in two steps. Let $S_n = X_1 + \cdots + X_n$ and define $\{X'_i\}_{i \geq 0}$ using

$$X'_i = \frac{1}{\sqrt{n}} S_{f(nt)},$$

where $f(nt)$ determines the indexing of $S_n$. Equation 2.4 is an approximation to the original Wiener process, and approximates the distribution law of the latter as the sample size gets large. Several alternative methods exist to simulate this process. Again, see Chapter 8 of (Lamberton and Lapeyre 1996) for a more complete discussion. Stochastic processes with jumps can be simulated using the above together with a Poisson process.

It has been shown that a path in the HJM scenario tree shown earlier is the realization of a stochastic process describing the forward rate. Using the Wiener process described above, the forward rate for an $m$-factor model can be simulated using

$$f(t + \Delta, T) - f(t, T) = \alpha(t, T; s_i)\Delta + \sum_{i=1}^{m} \sigma_i(t, T; s_i) W_i(t + \Delta) \sqrt{\Delta}, \quad (2.5)$$

where
for details on this implementation, see (Carr and Yang 1996).

2.2 Numerical issues
Monte Carlo simulation methods are widely applicable, but several caveats should be kept in mind when applying this approach in practice:

- The variance of the sampled variable remains usually very large, and the standard error decreases as $1/\sqrt{N}$, where $N$ is the number of samples. Two different approaches are used address this problem: variance reduction methods make use of effective sampling strategies, which are sometimes obtained at the expense of losing convergence criteria. The other aims at constructing low discrepancy sequences to speed up the convergence rate. For details, see (Broadie and Glasserman 1997; Broadie, Glasserman et al. 1997).

- The estimators obtained from Monte Carlo simulation are usually biased. This depends on the choice of sampling function.

- Finally, a point which is often overlooked in practice, the choice of random number generators should be made with care to fit the specific simulation problem. In particular, one should carefully select the seeds of the generators so that no cycling in the realization of the variables takes place. This yields redundancy in the generated sample, and affects convergence.

2.3 The Cholesky factorization
As seen earlier, this is an important step in simulating multi-variate Gaussian variables. Cholesky factorization is really a special case of LU decomposition for symmetric positive definite matrices. One starts with the $n \times n$ matrix $G$, symmetric positive definite, with elements given by $\sigma_{ij}$. The problem is to find a matrix

$$A^T = \begin{bmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ a_{21} & a_{22} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \ldots & \ldots & \ldots & a_{nn} \end{bmatrix}$$

such that $G = A^T A$. The simplest algorithm consists in applying a recursive transformation to the lower triangular part of $G$, column-wise, i.e., starting with the first column, and generate column $k$ given column $k-1$. The recursive definition for the diagonal elements is
and the off-diagonal lower triangular elements are given by

\[
a_{ij} = \frac{1}{a_{ii}} \left( \sigma_{ij} - \sum_{k=1}^{i-1} a_{ik} a_{jk} \right)^{1/2},
\]

for \( j = i + 1, \ldots, n \). All other elements are zero. The work necessary to construct this matrix is in \( O(n^3) \). Many variants of this algorithm exist, in particular by adjunction of a penalty matrix when it comes to handle sparsity of the original matrix.

### 2.4 The RiskMetrics software

RiskMetrics is software package written and maintained by J.P. Morgan. It consists of two parts. The first is a large database containing the historical evolution of significant market factors (interest, exchange rates, equity prices, etc.). In particular the database contains changes in volatility and correlations between factors normalized over daily time intervals.\(^1\)

The second part consist of a set of analytic tools referencing the database and used to compute market risk. This part contains an implementation of scenario generation using Monte Carlo methods, which is described briefly. For details, see (JPMorgan/Reuters 1996).

A lognormal distribution is assumed for the price of all underlying instruments and all market factors. Given \( P_0 \), the price/state of an instrument/factor at \( t = 0 \), and \( \sigma \), the one-day estimate for the volatility of this instrument/state, the evolution of the process as a function of time is assumed to be given by

\[
P_t = P_0 e^{\sigma \sqrt{t} X},
\]

where \( X \) is a standard Gaussian variable obtained by simulation. The example below (JPMorgan/Reuters 1996) follows the construction of a normally distributed random vector shown in the preceding section. The goal is to obtain the vector of future states of a standard \( 2 \times 1 \) Gaussian vector \( \bar{X} \) with a covariance matrix given by

\[
G = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.
\]

\(^1\)It should be noted that several problems exist with computing volatility from historical data. The variance of stock prices and interest rates has been empirically verified to be a non-stationary process, so large samples taken over time will not necessarily decrease the sampling errors (and in fact may actually increase it!).
The first step is to obtain

\[ \hat{Q} = (q_1, q_2)^T, \]  

(2.12)
a vector of standard Gaussian variables obtained from the first example. The second is to define

\[ \hat{X} = A\hat{Q} \]  

(2.13)
where

\[ A = \begin{bmatrix} 1 & \rho \sqrt{1 - \rho^2} \\ 0 & 1 \end{bmatrix} \]  

(2.14)
is the Cholesky decomposition of \( G \) (recall that \( G \) is symmetric positive definite). It can be verified that \( X \) is a standard Gaussian vector. To generate scenarios, one uses the future state equations given earlier,

\[ P^1_t = P^1_0 e^{\sigma_1 \sqrt{t} \hat{X}_1} \]
\[ P^2_t = P^2_0 e^{\sigma_2 \sqrt{t} \hat{X}_2} \]  

(2.15)
and repeats the process \( N \) times. As \( N \) gets large, the vector above approximates a lognormal distribution.

### 3.0 Scenario generation in RiskWatch

RiskWatch essentially provides a user interface for the manual generation of scenarios, extended to permit some limited analytic generation. A scenario in RiskWatch is defined by the following attributes:

- a weight (probability of occurrence),
- one (or more) variable(s) representing either instruments, positions or curves,
- a type of scenario. For all variables, one can choose between a constant factor perturbation (value multiplied by a constant), or a parallel shift perturbation (constant added to the value). For the curve variable only, two more types exist: a variable factor perturbation, in which the curve undergoes a parallel up or down shift, and a non-parallel shift, under which some proportional scaling is applied to all data points on the curve.

Five different scenario generation modes are implemented in the software:

- Manual. The scenario universe is entirely defined by the user.
- Bucket generation. A perturbation is applied to each point on the curve variable in an automatic fashion, one point at a time, each perturbation generating one scenario.
• Normal. A set of scenario with a normally distributed probability distribution function is generated.
• Monte Carlo generation: the approach developed in RiskMetrics and described earlier is applied to generate one-period scenarios of future prices (columns of the $D$ matrix).
• Multi-step Monte Carlo generation: As above, but with a user defined mean reversion coefficient applied to every new scenario curve.

Bucket and normal generation are only used for limited stress testing. The Monte Carlo modes are used in the optimization module.

References


