Letian Wang

HJM Model

HJM model is not a transitional model that bridges popular LIBOR market model with once popular short rate models, but an important framework that encompasses most of the interest rate models in the market. As the first multi-factor model, its functionality needs to be appreciated. In the beginning of this chapter, the model is introduced. It is followed by the demonstration of how to get from HJM to short rate models and LIBOR market models. HJM Principal Component analysis (PCA) is a historical estimation/calibration method to find HJM model parameters from historical data, as opposed to from current market quotes. It sheds light on historical estimation technique for LIBOR market models in the next chapter.

Contents

HJM Model

Heath, Jarrow, and Morton (1992) discovered the no-arbitrage condition between the instantaneous standard deviation and drifts of forward rates under risk-neutral measure Q, where the numeraire is the bank account $B(t)$. It is assumed that, **for a fixed maturity T**, the instantaneous forward rate $f(t,T)$ under Q follows:

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T) \cdot dW(t)$$
$$f(0,T) = f^M(0,T)$$

where operator $\cdot$ is the inner product of two vectors; $W = (W_1, \ldots, W_N)^T$ and $\sigma(t,T) = \left(\sigma_1(t,T), \ldots, \sigma_N(t,T)\right)^T$. If in addition $\sigma(t,T)$ does not depend on $f(t,T)$, it is known as the Gaussian HJM, where the instantaneous forward rates are normally distributed.

Note that for each fixed T there is one $f(t,T)$, so in total there are infinite many/continuum equations.

By definition, we have

$$f(t,T) = -\frac{\partial lnP(t,T)}{\partial T}$$
$$P(t,T) = exp\left\{-\int_t^T f(t,u)du\right\} = exp\{Q(t,T)\}$$

where

$$Q(t,T) = -\int_t^T f(t,u)du$$

Heath, Jarrow, and Morton (1992) claimed that under risk-neutral measure, the drift term $\alpha(t,T)$ can't be arbitrary but determined by the diffusion term $\sigma(t,T)$. Let's show this relationship.

First, take derivative on $Q(t,T)$,

$$dQ(t,T) = f(t,t)dt - \int_t^T df(t,u)du$$
$$= r(t)dt - \int_t^T \{\alpha(t,u)dt + \sigma(t,u) \cdot dW(t)\}du$$
$$= r(t)dt - \left(\int_t^T \alpha(t,u)du\right)dt - \left(\int_t^T \sigma(t,u)du\right) \cdot dW(t)$$

Let

$$\alpha^*(t,T) = \int_t^T \alpha(t,u)du \qquad (1)$$

$$\sigma^*(t,T) = \int_t^T \sigma(t,u)du \tag{2}$$

So $Q(t,T)$ can be re-written as

$$dQ(t,T) = r(t)dt - \alpha^*(t,T)dt - \sigma^*(t,T) \cdot dW(t)$$

Then

$$dP(t,T) = exp\{Q(t,T)\}\left\{dQ + \frac{1}{2}d\langle Q,Q\rangle\right\}$$

$$= P(t,T)\left\{\left(r(t) - \alpha^*(t,T) + \frac{1}{2}\sigma^*(t,T) \cdot \sigma^*(t,T)\right)dt - \sigma^*(t,T) \cdot dW(t)\right\}$$

Since the zero-coupon bond, as a tradable asset, its discounted price $P(t,T)/B(t)$ is a martingale under risk-neutral measure, or in other words, the drift term of $P(t,T)$ should be $r(t)P(t,T)$, we have

$$\alpha^*(t,T) = \frac{1}{2}\sigma^*(t,T) \cdot \sigma^*(t,T)$$

Take derivative w.r.t. T, it becomes

$$\alpha(t,T) = \sigma(t,T) \cdot \int_t^T \sigma(t,u)du = \sigma(t,T) \cdot \sigma^*(t,T) \tag{3}$$

In other words, **the drift term is completely determined by the volatility term**. This is the main result of Heath, Jarrow, and Morton (1992).

Therefore under risk-neutral world, the HJM model is

$$df(t,T) = \left\{\sigma(t,T) \cdot \int_t^T \sigma(t,u)du\right\}dt + \sigma(t,T) \cdot dW(t) \tag{4}$$

$$f(0,T) = f^M(0,T) \tag{5}$$

where $f^M(0,T)$, as exogenous inputs, ensuring that these models are automatically consistent with discount bond prices at time 0.

In addition, we have

$$f(t,T) = f(0,T) + \int_0^t \left\{\sigma(u,T) \cdot \int_u^T \sigma(u,s)ds\right\}du + \int_0^t \sigma(s,T) \cdot dW(s) \tag{6}$$

$$\frac{dP(t,T)}{P(t,T)} = r(t)dt - \left(\int_t^T \sigma(t,s)ds\right) \cdot dW(t) = r(t)dt - \sigma^*(t,T) \cdot dW(t) \quad (7)$$

$$P(t,T) = P(0,T) + \int_0^t P(s,T)r(s)ds - \int_0^t P(s,T)\sigma^*(s,T)dW(s) \quad (8)$$

$$r(t) = f(t,t) = f(0,t) + \int_0^t \left\{\sigma(u,t) \cdot \int_u^t \sigma(u,s)ds\right\}du + \int_0^t \sigma(s,t) \cdot dW(s) \quad (9)$$

From (9), the short rate process r(t) is not a Markov process in general.

From HJM to HW

HW (Chapter 9) is a special case of one factor HJM (N = 1),

$$df(t,T) = \left\{\sigma(t,T)\int_t^T \sigma(t,u)du\right\}dt + \sigma(t,T)dW(t)$$

and assume time-homogeneous exponential form

$$\sigma(t,T) = \sigma e^{-a(T-t)} \quad (10)$$

Then

$$\begin{aligned} r(t) &= f(0,t) + \int_0^t \left\{\sigma(u,t)\int_u^T \sigma(u,s)ds\right\}du + \int_0^t \sigma(s,t)dW(s) \\ &= f(0,t) + \int_0^t \left\{\sigma e^{-a(t-u)}\int_u^T \sigma e^{-a(s-u)}ds\right\}du + \int_0^t \sigma e^{-a(t-s)}dW(s) \\ &= f(0,t) + \frac{\sigma^2}{2a^2}(1-e^{-at})^2 + \sigma\int_0^t e^{-a(t-u)}dW(s) \\ &= r(0)e^{-at} + f(0,t) + \frac{\sigma^2}{2a^2}(1-e^{-at})^2 - f(0,0)e^{-at} + \sigma\int_0^t e^{-a(t-u)}dW(s) \end{aligned}$$

Compare it with the HW model (Chapter 9) and notice that $f(0,0) = r(0)$, we have

$$dr(t) = [\theta(t) - ar(t)]dt + \sigma dW(t)$$
$$\theta(t) = \frac{\partial f(0,t)}{\partial T} + af(0,t) + \frac{\sigma^2}{2a}(1-e^{-2at})$$

From HJM to G2++

G2++ is also a special case of HJM where there are two factors (N=2). It becomes

$$df(t,T) = \alpha(t,T)dt + \sigma_1(t,T)dW_1(t) + \sigma_2(t,T)dW_2(t)$$
$$dW_1(t)dW_2(t) = \rho dt$$

In reality, the two factors can be determined by principal components analysis (PCA): $W_1(t)$ stands for the change of slope, while $W_2(t)$ stands for the change of curvature.

Assume time-homogeneous exponential form

$$\sigma(t,T) = \sigma e^{-a(T-t)} + \eta e^{-b(T-t)} \tag{11}$$

Follow the same logic as in the previous section

$$r(t) = f(0,t) + r(0)e^{-at} + \left\{\frac{\sigma^2}{2a^2}(1-e^{-at})^2 - f(0,0)e^{-at}\right\} + \sigma\int_0^t e^{-a(t-u)}dW_1(s)$$
$$+r(0)e^{-bt} + \left\{\frac{\sigma^2}{2b^2}\left(1-e^{-bt}\right)^2 - f(0,0)e^{-bt}\right\} + \eta\int_0^t e^{-b(t-u)}dW_2(s)$$

Let

$$\varphi(t) = f(0,t) + r(0)e^{-at} + \left\{\frac{\sigma^2}{2a^2}(1-e^{-at})^2 - f(0,0)e^{-at}\right\}$$
$$+r(0)e^{-bt} + \left\{\frac{\sigma^2}{2b^2}\left(1-e^{-bt}\right)^2 - f(0,0)e^{-bt}\right\}$$
$$x(t) = \sigma\int_0^t e^{-a(t-u)}dW_1(s)$$
$$y(t) = \eta\int_0^t e^{-b(t-u)}dW_2(s)$$

Then

$$dx(t) = -ax(t)dt + \sigma dW_1(t), \quad x(0) = 0$$
$$dy(t) = -bx(t)dt + \eta dW_2(t), \quad y(0) = 0$$
$$dW_1(t)dW_2(t) = \rho dt$$

and it arrives the G2++ model under the risk-neutral measure

$$r(t) = x(t) + y(t) + \varphi(t)$$

From HJM to LMM

In this section we discuss the forward measure, caplet pricing, and LIBOR market model.

Consider the forward rate

$$F(t;S,T) = \frac{1}{\tau(S,T)}\left(\frac{P(t,S)}{P(t,T)} - 1\right)$$

Apply Ito lemma,

$$dF(t;S,T) = \frac{1}{\tau(S,T)}\left\{\frac{1}{P(t,T)}dP(t,S) + P(t,S)d\left(\frac{1}{P(t,T)}\right) + dP(t,S)d\left(\frac{1}{P(t,T)}\right)\right\} \quad (12)$$

Notice that from equation (7),

$$dP(t,S) = P(t,S)[r(t)dt - \sigma^*(t,S)\cdot dW(t)]$$

$$\begin{aligned} d\left(\frac{1}{P(t,T)}\right) &= -\frac{1}{P(t,T)^2}dP(t,T) + \frac{1}{P(t,T)^3}d\langle P(t,T), dP(t,T)\rangle \\ &= \frac{1}{P(t,T)}\{(-r(t) + \|\sigma^*(t,T)\|^2)dt + \sigma^*(t,T)\cdot dW(t)\} \end{aligned}$$

so that equation (12) can be re-written as

$$\begin{aligned} dF(t;S,T) &= \frac{1}{\tau(S,T)}\left\{\frac{1}{P(t,T)}dP(t,S) + P(t,S)d\left(\frac{1}{P(t,T)}\right) + dP(t,S)d\left(\frac{1}{P(t,T)}\right)\right\} \\ &= \frac{1}{\tau(S,T)}\frac{P(t,S)}{P(t,T)}\{\left(|\sigma^*(t,T)|^2 - \sigma^*(t,S)\cdot\sigma^*(t,T)\right)dt + \left(\sigma^*(t,T) - \sigma^*(t,S)\right)\cdot dW(t)\} \\ &= \left[F(t;S,T) + \frac{1}{\tau(S,T)}\right][\sigma^*(t,T) - \sigma^*(t,S)]\cdot[\sigma^*(t,T)dt + dW(t)] \end{aligned}$$

That is,

$$dF(t;S,T) = \left[F(t;S,T) + \frac{1}{\tau(S,T)}\right][\sigma^*(t,T) - \sigma^*(t,S)]\cdot[\sigma^*(t,T)dt + dW(t)] \quad (12)$$

Now define the shifted LIBOR forward rate as

$$F^*(t;S,T) = F(t;S,T) + \frac{1}{\tau(S,T)} \quad (13)$$

Then it follows lognormal process

$$dF^*(t;S,T) = F^*(t;S,T)\left[\int_S^T \sigma(t,u)du\right]\cdot[\sigma^*(t,T)dt + dW(t)] \qquad (14)$$

We know under the T-forward measure, the (shifted) LIBOR forward rate is a martingale. To **change the measure from risk-neutral to T-forward measure**, use Girsanov Theorem (Wikipedia: Girsanov Theorem)

$$X(t) = -\int_0^t \left(\int_s^T \sigma(s,u)du\right) dW(s) = -\int_0^t \sigma^*(s,T)dW(s)$$
$$dX(t) = -\left(\int_t^T \sigma(t,u)du\right)dW(t) = -\sigma^*(t,T)dW(t)$$
$$\varepsilon(X(t)) = exp\left\{X(t) - \frac{1}{2}[X(t)]\right\}$$
$$W^T(t) = W(t) - X(t)$$

$$\boldsymbol{dW^T(t)} = dW(t) - [dW(t), dX(t)] = \boldsymbol{dW(t) + \sigma^*(t,T)dt} \qquad (15)$$

So that under T-forward measure

$$dF^*(t;S,T) = F^*(t;S,T)\left[\int_S^T \sigma(t,u)du\right]\cdot W^T(t) \qquad (16)$$

In addition, from equation (4), under T-forward measure, the instantaneous forward rate is a martingale, or

$$df(t,T) = \sigma(t,T)\cdot dW^T(t) \qquad (17)$$

This can be seen from

$$df(t,T) = \sigma(t,T)\cdot\sigma^*(t,T)dt + \sigma(t,T)\cdot dW(t)$$
$$= \sigma(t,T)\cdot\sigma^*(t,T)dt + \sigma(t,T)\cdot[dW^T(t) - \sigma^*(t,T)dt]$$
$$= \sigma(t,T)\cdot dW^T(t)$$

In other words, equation (16) shows that the expectation hypothesis holds under the forward measure, or

$$E^{Q^T}[r(T)|\mathrm{F}_t] = f(t,T)$$

Caplet/Floorlet on rate $F(S;S,T)$ with strike K can be re-defined on rate $F^*(S;S,T)$ with strike $K+\frac{1}{\tau(S,T)}$. The Black volatility is

$$\int_0^S \left\| \int_S^T \sigma(t,u)du \right\|^2 dt \tag{18}$$

Anderson and Piterbarg (2010) pointed out that Log-Normal (proportional volatility) HJM

$$\sigma(t,T,f(t,T)) = \sigma(t,T)f(t,T) \tag{19}$$

Leads to log-normally distributed instantaneous forward rate under T-forward measure

$$df(t,T) = \sigma(t,T)\cdot f(t,T)dW^T(t)$$

but the forward rates will explode in finite time to infinity with non-zero probability. In order to fix this drawback, it comes the LIBOR market model (Chapter 11).

Actually, LIBOR market model can be viewed as a special case of HJM model. Suppose there exists an deterministic function $\hat{\sigma}(t,T)$ such that

$$\int_S^T \sigma(t,u)du = \frac{F(t;S,T)}{F(t;S,T)+\frac{1}{\tau(S,T)}}\hat{\sigma}(t,T)$$

Then from equation (12),

$$dF(t;S,T) = \hat{\sigma}(t,T)F(t;S,T)dW^T(t)$$

HJM PCA

In this section it first introduces the historical estimation /calibration for HJM. Then it discusses the Principal Component Analysis (PCA) for HJM.

Firstly, one important assumption about historical estimation is time homogeneity. Consider Gaussian HJM model in equation (4) and (5), time homogeneity means that the volatility parameter $\sigma(t,T)$ depends only on the remaining time to maturity, or

$$\sigma(t,T) = \sigma(T-t)$$

so that equation (4) can be rewritten as

$$\Delta f(t,T) = \mu(T-t)\Delta t + \sigma(T-t)\cdot \Delta W(t)$$

It enables us to estimate σ (hence drift) from historical data. For example, if we want to estimate 3-month volatility

$$\sigma(t, T = t + 3M) = \sigma(T-t) = \sigma(3M)$$

On 2011-Jun-01, we record the change of instantaneous forward rate that ends on 2011-August-31; on 2011-Jun-02, we move forward and record the change of instantaneous forward rate that ends on 2011-Sep-01; on 2011-Jun-03, we record the historical move of instantaneous forward rate that ends on 2011-Sep-02, and so on. Technically, we move from one forward rate equation to another forward rate equation (remember there are infinite many of them). Since the time to maturity is kept at 3 month, in this way we obtain a series of historical observation on the very same $\sigma(3M)$. This is why we need time homogeneity.

The historical estimation can be done on instantaneous forward rate, or on more observable continuous zero rate. Recall that continuously-compounded spot interest rate is defined by

$$R(t,T) = -\frac{\ln P(t,T)}{\tau(t,T)}$$

Then apply Ito lemma on equation (7),

$$\begin{aligned} dR(t,T) &= -d\left(\frac{1}{\tau(t,T)}\right)\ln P(t,T) - \frac{1}{\tau(t,T)}\left\{\frac{1}{P(t,T)}dP(t,T) - \frac{1}{2P(t,T)^2}d\langle P(t,T),P(t,T)\rangle\right\} \\ &= -\frac{\ln P(t,T)}{(T-t)^2}dt - \frac{1}{\tau(t,T)}\left\{\left(r(t) - \frac{1}{2}\|\sigma^*(t,T)\|^2\right)dt - \sigma^*(t,T)\cdot dW(t)\right\} \\ &= \frac{1}{\tau(t,T)}\left\{\left(R(t,T) - r(t) + \frac{1}{2}\|\sigma^*(t,T)\|^2\right)dt + \sigma^*(t,T)\cdot dW(t)\right\} \end{aligned}$$

which yields

$$dR(t,T) == \frac{1}{\tau(t,T)}\left\{\left(R(t,T) - r(t) + \frac{1}{2}\|\sigma^*(t,T)\|^2\right)dt + \sigma^*(t,T)\cdot dW(t)\right\} \quad (20)$$

In the rest of this section, we consider general processes

$$dX_i(t) = \mu_i dt + \sum_{j=1}^{N}\sigma_{ij}dW_j(t), \quad i = 1, \dots, m \qquad (21)$$

where X can be instantaneous forward rate (then it becomes equation (4)), or continuously-compounded spot rate (then it becomes equation (20)). In equation (21), time T (or T-t) is replaced by subscript i, to indicate explicitly the discrete tenors (e.g. i=1 standards for 3-month time to maturity, i=2 standards for 6-month time to maturity, etc.). That is,

$$\sigma_{ij} = \sigma_j(t, t + \tau_i)$$

Secondly, in order to understand better the mechanism of PCA and later the LIBOR market model, it's helpful to re-write equations (21) in the following format. Define

$$\sigma_i = \sqrt{\sum_{j=1}^{d} \sigma_{ij}^2}$$

and define processes

$$Z_i(t) = \sum_{j=1}^{N} \frac{\sigma_{ij}}{\sigma_i} W_j(t), \quad i = 1, \ldots, m$$

$$dZ_i(t) dZ_i(t) = \sum_{j=1}^{N} \frac{\sigma_{ij}^2}{\sigma_i^2} dt = dt$$

$$dZ_i(t) dZ_k(t) = \sum_{j=1}^{N} \frac{\sigma_{ij}\sigma_{kj}}{\sigma_i \sigma_k} dt = \rho_{ik} dt$$

where

$$\rho_{ik} = \frac{1}{\sigma_i \sigma_k} \sum_{j=1}^{N} \sigma_{ij}\sigma_{kj}$$

Then the equations (21) can be re-written in terms of the Brownian motion $Z_i(t)$ as

$$dX_i(t) = \mu_i dt + \sigma_i dZ_i(t), \quad i = 1, \ldots, m \qquad (22)$$

$$dZ_i(t) dZ_k(t) = \rho_{ik} dt \qquad (23)$$

As explained before, the volatility σ_i is constant here because of time-homogeneity. Also the drift term μ_i is completely determined by σ_i from no-arbitrage argument.

Therefore one can estimate the variance-covariance matrix and obtain $\sigma_i, \rho_{i,k}$ from the historical movements of $\Delta X_i(t)$. The estimation process is standard (sample variance and covariance) hence is omitted here. If one's task is to estimate/calibrate the HJM model, the task is achieved by now.

Often in practice, we go one step further and simulate the estimated/calibrated HJM models. The Monte Carlo simulation of HJM is commonly seen in pricing exotic products (such as swaptions), or in counterparty credit risk calculations (such as PFE calculation). It is suffered by "curse of dimensionality", in this case dimension is m.

In the next step we reduce the dimension via PCA. Usually first three PCAs are picked, standing for three main yield curve movements: parallel shifts, twists (steepen/flatten), and curvature (butterfly). The logic is simple: we have m variables that are correlated in equation (22), and we want to find m uncorrelated variables in equation (21). Use the jargons of PCA, we want to

$$\{dZ_1, \dots, dZ_m\} \overset{PCA}{\Longrightarrow} \{dW_1, \dots, dW_m\}$$

and then pick the largest 3 PCAs by setting $N = 3$ in equation (21). PCA is also a standard statistics practice hence is neglected here. The method is briefly introduced in the appendix.

In the next chapter LMM model will be discussed. In terms of calibration, yield correlation matrix can be obtained from either historical data or from the prices of caps and swaptions. Readers may have already figured out how to calibrate LMM through historical estimation, yes, it is quite similar to what we have seen in this section. How to calibrate to caps and swaptions is the main topic of the next chapter.

Appendix 1 Principal Component Analysis

Denote $X \in R^{T\times n}$ the matrix of data time series, where T is the number of data points and n is the number of variables.

$$X = [x_1, x_2, \dots, x_n] \in R^{T\times n}$$

We assume each column of X has zero mean. If in addition they are normalized, then the correlation matrix

$$C = \frac{1}{T} X^T X \in R^{n\times n}$$

Denote its eigenvalues by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ and their corresponding eigenvectors $w_1, w_2, \dots, w_n$, and

$$Cw_i = \lambda_i w_i, C = W\Lambda W^{-1}$$
$$\Lambda = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \in R^{n\times n}$$
$$W = [w_1, w_2, \dots, w_n] \in R^{n\times n}$$

so that Λ is the diagonal matrix of eigenvalues and W is the orthogonal matrix of eigenvectors ($W^T = W^{-1}$) because C is symmetric.

The matrix of principal components of C is a $T \times n$ matrix P defined by

$$P = [p_1, p_2, \dots, p_n] = XW$$

The mth principal component of C is defined as the mth column of P,

$$p_m = w_{m1}x_1 + w_{m2}x_2 + \cdots + w_{mn}x_n$$

where $w_m = [w_{m1}, w_{m2}, \dots, w_{mn}]^T$ is the eigenvector corresponding to λ_m. Thus p_1 belongs to the first and the largest eigenvalue λ_1, p_2 belongs to the first and the largest eigenvalue λ_2, and so on. Each principal component is a time series, i.e., a $T \times 1$ vector.

The covariance matrix of the principal components is Λ, or

$$\frac{1}{T}P^T P = \frac{1}{T}W^T X^T X W = W^T C W = W^{-1} W \Lambda = \Lambda$$

So that the principal components are uncorrelated, the variance of the mth principal component is λ_m.

Now

$$X = PW^{-1} = PW^T$$

So that we can write each of the original variables as a linear combination of the principal components, as

$$x_i = w_{1i}p_1 + w_{2i}p_2 + \cdots + w_{ni}p_n$$
$$\approx w_{1i}p_1 + w_{2i}p_2 + w_{3i}p_3$$

Appendix 2 Taylor Expansion and Ito Lemma

1-D Taylor expansion

$$f(x) = f(x_0) + \frac{\partial f}{\partial x}(x - x_0) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x - x_0)^2 + \cdots$$

2-D Taylor expansion

$$f(x,y) = f(x_0,y_0) + \left[(x - x_0)\frac{\partial}{\partial x} + (y - y_0)\frac{\partial}{\partial y}\right]f$$
$$+ \frac{1}{2}\left[(x - x_0)\frac{\partial}{\partial x} + (y - y_0)\frac{\partial}{\partial y}\right]^2 f + \cdots$$

1-D Ito lemma

$$df = \frac{\partial f}{\partial x} dx + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} d\langle x, x\rangle$$

2-D Ito lemma

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{1}{2}\left\{\frac{\partial^2 f}{\partial x^2} d\langle x, x\rangle + 2\frac{\partial^2 f}{\partial x \partial y} d\langle x, y\rangle + \frac{\partial^2 f}{\partial y^2} d\langle y, y\rangle\right\}$$

In particular,

$$d(xy) = ydx + xdy + d\langle x, y\rangle$$

Reference

[1] Anderson, L. and Piterbarg, V. (2010). Interest Rate Modeling. Volume 1: Foundations and Vanilla Models. Atlantic Financial Press.

[2] Anderson, L. and Piterbarg, V. (2010). Interest Rate Modeling. Volume 2: Term Structure Models. Atlantic Financial Press.

[3] Brace, A. and M. Musiela (1994). A multi-factor Gauss Markov implementation of Heath, Jarrow and Morton. *Mathematical Finance*, 4, pp. 259-283.

[4] Brace, A. and M. Musiela (1997). Swap derivatives in a Gaussian HJM framework. In M. Dempster and S. Pliska (Eds.). *Mathematics of Derivatives Securities*, pp. 336-368. Cambridge University Press.

[5] Brigo, D. and Mercurio, F (2006). Interest rate models: theory and practice: with smile, inflation, and credit. Springer Verlag.

[6] Duffie, D. and Kan, R. (1996). A yield-factor model of interest rates. *Mathematical finance*, 6(4).

[7] Filipovic, D. (2009). Term-structure models: A graduate course, Springer Verlag.

[8] Heath, D. and Jarrow, R. and Morton, A.(1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60(1), pp. 77-105.

[9] Jamshidian, F. and Zhu, Y. (1997). Scenario Simulation: Theory and methodology. *Finance and Stochastics*. 1, pp. 43-67.

[10] Shreve, S.E. (2004). Stochastic calculus for finance. Springer

Header Image: Downtown Montreal from Chalet du Mont-Royal