Quantitative finance

Interviews preparation

Vivien Tisserand

Abstract

This is a summary of interview questions that I found and digressed on in quantitative finance. It is a mixed of applied mathematics and computer science.
I am not found of brainteasers, they are a poor way to assess for a candidate's hability to be an asset for the teams. This work smoothly transitioned to a sort of *vademecum* in applied mathematics : through several questions, it goes through several techniques that are easy to forget with time. I myself refer to it quite often when I forget about the way to solve an arithmetico-geometric sequence or the general solution of a second-order differential equation.

Contents

Chapter 1

Statistics

1.1 Estimating the support of an uniform law

Suppose that we have $x_1, \dots, x_n$ observations from an uniform law $X \sim \mathcal{U}[0, \theta]$, where θ is an unknown parameter that we want to estimate. Give at least two estimators for θ and compare them.

- **Method of moments:**

  Having a look at the first order moment, it appears that $\mathbb{E}[X] = \dfrac{\theta}{2}$. Taking the empirical counter-party of this theoretical quantity, we have $\hat{\theta}^{\mathrm{MM}} = \dfrac{2}{n} \sum_{i=1}^{n} x_i$.
  By applying the strong law of large numbers and the continuous mapping theorem, $\hat{\theta}^{\mathrm{MM}} \xrightarrow{a.s.} \theta$. Thus this estimator is consistent.
  We want asymptotic results on the convergence of this estimator. Before using the CLT, we have to check for the existence of a second-order moment.

$$
\begin{aligned}
\mathbb{E}[X^2] &= \int_{\mathbb{R}} x^2 f(x) \,\mathrm{d}x \\
&= \int_0^{\theta} x^2 \frac{1}{\theta} \,\mathrm{d}x \\
&= \left[\frac{1}{3\theta} x^3 \right]_0^{\theta} \\
&= \frac{\theta^2}{3} < +\infty
\end{aligned}
$$

  Thus, we have $\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \dfrac{\theta^2}{12}$. So, $\mathbb{V}[2X_1] = \dfrac{\theta^2}{3}$.

  By applying the central limit theorem, we have :

$$
\sqrt{n}(\hat{\theta}^{\mathrm{MM}} - \theta) \xrightarrow{(d)} \mathcal{N}\left(0, \frac{\theta^2}{3}\right).
$$

  We have to evaluate the risk of this estimator, that we write as the sum of the squared bias and the variance :

  $\mathrm{MSE}(\hat{\theta}^{\mathrm{MM}}) = \mathbb{E}[(\hat{\theta}^{\mathrm{MM}} - \theta)^2] = \mathbb{E}[(\hat{\theta}^{\mathrm{MM}} - \mathbb{E}[\hat{\theta}^{\mathrm{MM}}])^2] + \mathbb{E}[\hat{\theta}^{\mathrm{MM}} - \theta]^2 = \mathbb{V}[\hat{\theta}^{\mathrm{MM}}] + (\mathbb{E}[\hat{\theta}^{\mathrm{MM}}] - \theta)^2.$

We have $\mathbb{E}[\hat{\theta}^{\mathrm{MM}}] = 0$ and $\mathbb{V}[\hat{\theta}^{\mathrm{MM}}] = \dfrac{1}{n^2} n \mathbb{V}[2X_1] = \dfrac{\theta^2}{3n}$.

Thus,

$$\mathrm{MSE}(\hat{\theta}^{\mathrm{MM}}) = \frac{\theta^2}{3n}.$$

- **Maximum likelihood:**

  Let's write the likelihood of this model:

$$\begin{aligned}
L((X_1, \ldots, X_n), \theta) &= \prod_{i=1}^{n} f_X(X_i) \\
&= \prod_{i=1}^{n} \frac{1}{\theta} \mathbb{1}_{[0,\theta]}(X_i) \\
&= \frac{1}{\theta^n} \prod_{i=1}^{n} \mathbb{1}_{[0,\theta]}(X_i).
\end{aligned}$$

  And this function is maximized by choosing the smallest θ such that all of the X_i lie in $[0, \theta]$, that is $\hat{\theta}^{\mathrm{MLE}} = \max_{1 \leq i \leq n} X_i$.

  To check the consistency of this estimator, we will have a look at its convergence (in probability). Let $\theta \in \Theta$ and $\varepsilon > 0$:

$$\begin{aligned}
\mathbb{P}_\theta(|\hat{\theta}^{\mathrm{MLE}} - \theta| \geq \varepsilon) &= \mathbb{P}_\theta(\hat{\theta}^{\mathrm{MLE}} \geq \theta + \varepsilon) + \mathbb{P}_\theta(\hat{\theta}^{\mathrm{MLE}} \leq \theta - \varepsilon) \\
&= 0 + \mathbb{P}_\theta(\max_{1 \leq i \leq n} X_i \leq \theta - \varepsilon) \\
&= \prod_{i=1}^{n} \mathbb{P}_\theta(X_i \leq \theta - \varepsilon) \\
&= \left(1 - \frac{\varepsilon}{\theta}\right)^n \underset{n \to +\infty}{\longrightarrow} 0.
\end{aligned}$$

  Thus, $\hat{\theta}^{\mathrm{MLE}} \xrightarrow{\mathbb{P}} \theta$: this estimator is consistent.

  In order to estimate the risk of this estimator, we have to look at the law that the maximum of n independent uniform laws follows. This is done by looking at the cumulative distribution function. Let $x \in [0, \theta]$:

$$\begin{aligned}
\mathbb{P}_\theta(X_{(n)} \leq x) &= \mathbb{P}_\theta\left(\bigcap_{i=1}^{n} X_i \leq x\right) \\
&= \prod_{i=1}^{n} \mathbb{P}_\theta(X_i \leq x) \\
&= \left(\frac{x}{\theta}\right)^n.
\end{aligned}$$

Thus,

$$F_{X_{(n)}} = \begin{cases} 0 & \text{if } x < 0 \\ \left(\dfrac{x}{\theta}\right)^n & \text{if } 0 \leq x \leq \theta \\ 1 & \text{if } x > \theta \end{cases}$$

This cdf as smooth as we need to take its derivative: that will be the density we were looking for:

$$f_{X_{(n)}}(x) = n\frac{x^{n-1}}{\theta^n}\mathbb{1}_{[0,\theta]}(x)$$

Let's compute the bias and the variance.

$$\mathbb{E}[\hat{\theta}^{\mathrm{MLE}}] = \int_{\mathbb{R}} x f_{X_{(n)}}(x)\,\mathrm{d}x = \int_0^\theta \frac{n}{\theta^n}x^n\,\mathrm{d}x = \frac{n}{\theta^n}\left[\frac{x^{n+1}}{n+1}\right]_0^\theta = \frac{n}{n+1}\theta.$$

Then, the bias is : $B(\hat{\theta}^{\mathrm{MLE}}) = \frac{n}{n+1}\theta - \theta = -\frac{1}{n+1}\theta \neq 0$. We can introduce a corrected estimator that we will consider too : $\hat{\theta}^{\mathrm{MLE}}_{\mathrm{corr}} = \frac{n+1}{n}\hat{\theta}^{\mathrm{MLE}}$, such that $\mathbb{E}[\hat{\theta}^{\mathrm{MLE}}_{\mathrm{corr}}] = \theta$: an unbiased estimator.

Then, we have

$$\mathbb{E}[(\hat{\theta}^{\mathrm{MLE}})^2] = \int_{\mathbb{R}} x^2 f_{X_{(n)}}(x)\,\mathrm{d}x = \int_0^\theta \frac{n}{\theta^n}x^{n+1}\,\mathrm{d}x = \frac{n}{\theta^n}\left[\frac{x^{n+2}}{n+2}\right]_0^\theta = \frac{n}{n+2}\theta^2.$$

And

$$\mathrm{MSE}(\hat{\theta}^{\mathrm{MLE}}) = \mathbb{E}[(\hat{\theta}^{\mathrm{MLE}} - \theta)^2] = \mathbb{E}[(\hat{\theta}^{\mathrm{MLE}})^2] - 2\theta\mathbb{E}[\hat{\theta}^{\mathrm{MLE}}] + \theta^2$$

Thus,

$$\mathrm{MSE}(\hat{\theta}^{\mathrm{MLE}}) = \frac{n}{n+2}\theta^2 - 2\frac{n}{n+1}\theta^2 + \theta^2 = \frac{2\theta^2}{(n+1)(n+2)}.$$

And

$$\mathrm{MSE}(\hat{\theta}^{\mathrm{MLE}}_{\mathrm{corr}}) = \left(\frac{n+1}{n}\right)^2\mathbb{E}[(\hat{\theta}^{\mathrm{MLE}})^2] - 2\frac{n+1}{n}\theta\mathbb{E}[(\hat{\theta}^{\mathrm{MLE}} + \theta^2 = \frac{\theta^2}{n(n+1)}.$$

- **Maximum a posteriori:**

  We write the likelihood of the model in terms of θ :

$$L((X_1,\ldots,X_n),\theta) = \frac{1}{\theta^n}\prod_{i=1}^n \mathbb{1}_{[0,\theta]}(X_i) = \frac{1}{\theta^n}\mathbb{1}_{[X_{(n)},=\infty[}(\theta).$$

Remark : the set such that $L(.,\theta) > 0$ is $[0,\theta]$: it depends on θ, thus the model is not regular. Keep that in mind when dealing with Fisher information for instance.

1. **Flat prior:**

   We apply the definition for a Bayesian estimator with a prior density π_0 :

$$\begin{aligned}\hat{\theta}^{\mathrm{B}} &= \frac{\displaystyle\int_{\Theta} \theta L(x,\theta)\pi_0(\theta)\,\mathrm{d}\lambda(\theta)}{\displaystyle\int_{\Theta} L(x,\theta)\pi_0(\theta)\,\mathrm{d}\lambda(\theta)} \\ &= \frac{\displaystyle\int_{X_{(n)}}^{+\infty} \theta^{-n+1}\,\mathrm{d}\theta}{\displaystyle\int_{X_{(n)}}^{+\infty} \theta^{-n}\,\mathrm{d}\theta} \\ &= \frac{n-1}{n-2} X_{(n)}.\end{aligned}$$

   Bias and MSE are not computed there for sanity reasons.

2. **Jeffreys prior:**

   The density function of this prior is proportional to the squareroot of the determinant of the Fisher information matrix.

   Thus we need to compute this quantity for this model, with n observations (as it is not regular, $I_n \neq nI_1$) :

$$I_n(\theta) = \mathbb{E}\left[\frac{\partial \log L_n(\theta)}{\partial \theta}^2\right]$$

   We have $I_n(\theta) = \mathbb{E}[(-n/\theta)^2] = \dfrac{n^2}{\theta^2}$

   (If we had taken the expectancy of the second-order derivative of the log-likelihood, we would not have had the same result has the model is not regular.)

   This gives us the noninformative prior (Jeffreys) : $\pi_0(\theta) \propto \theta^{-1}$.

Chapter 2

Options mathematics

2.1 Dupire formula for local volatility

Explain the motivation behind local volatility and prove Dupire formula.

The Black-Scholes model is really convenient as it has few parameters, yields closed-form vanilla prices and is widely known. However, the assumption of constant volatility doesn't match the observed market prices.
Indeed, looking at the implied volatility surface we observe smile / skew / smirk accross all asset classes. In equities, the volatility smile appeared after the 1987 crisis and reflected this premium buyers were ready to pay to hedge against the downside risk.
In order to build new pricing models that reflected this stylized fact, a maturity and strike dependent volatility was introduced: going from a constant σ accros all instruments to $\sigma(t, S_t)$.

What is now known as the Dupire formula is the following expression for such a volatility function:

$$\sigma(t, S_t) = \frac{\dfrac{\partial C}{\partial T} + (r-q)K\dfrac{\partial C}{\partial K} + qC}{\dfrac{1}{2}K^2\dfrac{\partial^2 C}{\partial K^2}}.$$

We assume the underlying follows a Geometric brownian motion dynamic: $dS_t = \mu S_t dt + \sigma^2 S_t dW_t$, with $\mu = r - q$. We also introduce the discount factor between time t and maturity T: $D(t,T) = \exp\left(-\int_t^T r(s)ds\right)$.
The call option price can then be written as $C(K,T) = D(t,T)\mathbb{E}_{\mathbb{Q}[(S_T-K)^+]}$.

We are interested in the probability density of the underlying at maturity: $p(S,t)$. Its variations are governed by the **Fokker-Planck equation**:

$$\frac{\partial}{\partial t}p(S,t) = -\frac{\partial}{\partial S}(\mu S p(S,t)) + \frac{1}{2}\frac{\partial^2}{\partial \sigma^2}(\sigma^2 S^2 p(S,t)).$$

Let's compute the theta of a call option:

$$\frac{\partial C}{\partial T} = \frac{\partial D(t,T)}{\partial T}\int_K^{+\infty}(S-K)p(S,T-t)dS + D(t,T)\int_K^{+\infty}(S-K)\frac{\partial p(S,T-t)}{\partial T}dS.$$

Plugging in we get:

$$
\begin{aligned}
\Theta + rC &= D(t,T) \int_K^{+\infty} (S-K) \left[-\frac{\partial}{\partial S}(\mu S p(S,t)) + \frac{1}{2}\frac{\partial^2}{\partial \sigma^2}(\sigma^2 S^2 p(S,t)) \right] \\
&= D(t,T) \left(-\mu I_1 + \frac{1}{2} I_2 \right).
\end{aligned}
$$

We consider the first and second order derivatives with regards to the strike. Is is known that they respectively are equal to the cumulative distribution function above the strike and the probability density at maturity (the latter being the **Breeden-Litzenberger formula**). To know what quantities we should further consider, we apply integration by parts to I_1 and I_2, with the goal to get rid of integrands and fuzzy terms.

$$
\begin{aligned}
I_1 &= \int_K^{+\infty} (S-K) \frac{\partial}{\partial S}(S p(S,t)) \\
&= [(S-K)p(S,t)]_{S=K}^{S=+\infty} - \int_K^{+\infty} S p(S,t) dS \\
&= - \int_K^{+\infty} S p(S,t) dS.
\end{aligned}
$$

To explicit this last line, let's rewrite the call price as $\mathrm{C} = \mathrm{S}e^{-qT}(d_1) - Ke^{-rT}(d_2)$ such that $\displaystyle\int_K^{+\infty} S p(S,t) dS = \frac{1}{D(t,T)} \left(C - K \frac{\partial C}{\partial K} \right)$.

Then,

$$
\begin{aligned}
I_2 &= \int_K^{+\infty} (S-K) \frac{\partial^2}{\partial \sigma^2}(\sigma^2 S^2 p(S,t)) dS \\
&= \left[(S-K) \frac{\partial}{\partial \sigma}(\sigma^2 S^2 p(S,t)) \right]_{S=K}^{S=+\infty} - \frac{\partial}{\partial \sigma}(\sigma^2 S^2 p(S,t)) dS \\
&= - \left[(\sigma^2 S^2 p(S,t) \right]_{S=K}^{S=+\infty} \\
&= \sigma^2 K^2 p(S,t) \\
&= \sigma^2 K^2 \frac{1}{D(t,T)} \frac{\partial^2 C}{\partial K^2}.
\end{aligned}
$$

Going back to the theta derivation, we have

$$
\frac{\partial C}{\partial T} + rC = C - K \frac{\partial C}{\partial K} + \sigma^2 K^2 \frac{\partial^2 C}{\partial K^2}.
$$

Rearranging the terms, we get the Dupire formula.

There also exist a probabilistic derivation of this formula, applying Itô to the payoff $(S_T - K)^+$ and taking the expectation.