From Martingales in Finance to Quantization for pricing

Giorgia Callegaro

Università di Padova

Workshop on Martingales in Finance and Physics
ICTP, 24 May 2019.

Mostly based on recent papers with **Lucio Fiorin** and **Martino Grasselli**.

Martingales in Finance: why?

- **[Arbitrage]** Intuition: a possibility of a riskless profit. An arbitrage is an investment strategy whose cost today is non positive, whose (portfolio) value tomorrow is non-negative and strictly positive with positive probability (recall today's first talk).
- **[Viability]** A market is viable if there is no arbitrage opportunity: we ♡ AOA.
- **[Key theorem]** The market is viable if and only if there exists a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that the discounted asset prices are $\mathbb{Q}$-martingales.
  $\mathbb{P}$: real world measure
  $\mathbb{Q}$: risk neutral probability

Pricing and hedging in a nutshell

- **[Call and Put]** A Call **option** gives the holder the opportunity to buy an underlying asset X, at a fixed time T and at a specified cost (strike) $K > 0$: its value is

$$F_T^{\mathscr{C}} = (X_T - K)_+$$

The Put option analogously gives the holder the right to sell:

$$F_T^{\mathscr{P}} = (K - X_T)_+$$

- **[Pricing]** The price at time t of a European option, whose payoff is $F_T = f(X_T)$ is

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)} f(X_T) | \mathscr{F}_t\right]$$

where $(\mathscr{F}_t)_{t\in[0,T]}$ is the available filtration.

- **[Hedging]** An investment strategy whose portfolio's value coincides (replicates) at any time with the option's value.

What if we discretize X_T?

In case when the random variable X_T reduces to a finite set of points, the expectation (price) is computed as a finite sum.
Quantization: approximating a signal (random variable) admitting a continuum of possible values, by a signal that takes values in a discrete set

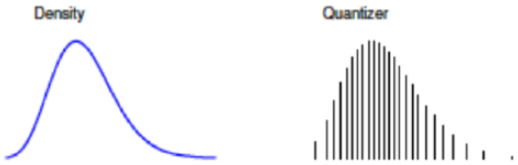


Figure: Picture taken from Mc Walter et al [9]

Quantization: a brief history

- **[Birth]** Back to the 50's, to optimize signals' transmission
- **[Two worlds]**
  - *Vector* quantization $\leftrightsquigarrow$ random variables
  - *Functional quantization* $\leftrightsquigarrow$ stochastic processes
- **[Applications]** Information theory, cluster analysis, pattern and speech recognition, numerical integration and probability
- **[How?]** Numerical procedures mostly based on stochastic optimization algorithms $\rightsquigarrow$ very time consuming.

Today's menu: discretize random variables and stochastic processes in a fast and efficient way via recursive marginal quantization.

Vector quantization: some Math

Given an $\mathbb{R}^d$-valued random variable X on $(\Omega, \mathscr{A}, \mathbb{P})$ (or $\mathbb{Q}$), $X \in L^r$, N-quantizing X on a grid $\Gamma = (x_1, \ldots, x_N)$ consists in projecting X on Γ. In order to univocally define the projection function, we need to specify a partition of $\mathbb{R}^d$, $(C_i)_{1 \le i \le N}$, so that

$$\mathrm{Proj}_\Gamma(X) = \sum_{i=1}^{N} x_i \mathbb{1}_{C_i}(X)$$

- The induced L^r error

$$||X - \mathrm{Proj}_\Gamma(X)||_r = \mathbb{E}\left[\min_{1 \le i \le N} |X - x_i|^r\right]^{1/r}$$

is called the L^r−mean quantization error.

N.B. For a complete background on optimal quantization: Graf and Luschgy [7]

Quadratic optimal quantization ($r = 2$)

Let us focus, from now on, on $r = 2$!

How do we choose the N points in Γ?
By minimizing the L^2 error!

- A grid $\Gamma^\star$ minimizing the L^2- quantization error over all the grids with size at most N is the optimal quadratic quantizer.
- The projection of X on $\Gamma^\star$, $\mathrm{Proj}_{\Gamma^\star}(X)$, or $\widehat{X}^{\Gamma^\star}$, or $\widehat{X}$ for simplicity, is called the quantization of X and the associated partition

$$C_i(\Gamma^\star) \subset \left\{ \xi \in \mathbb{R}^d : |\xi - x_i| = \min_{1 \leq j \leq N} |\xi - x_j| \right\}$$

is called the Voronoï partition, or tessellation induced by $\Gamma^\star$.
$\mathrm{Proj}_{\Gamma^\star}(X)$ is defined as the closest neighbor projection on $\Gamma^\star$.

Vector quantization: example ($N = 50$)

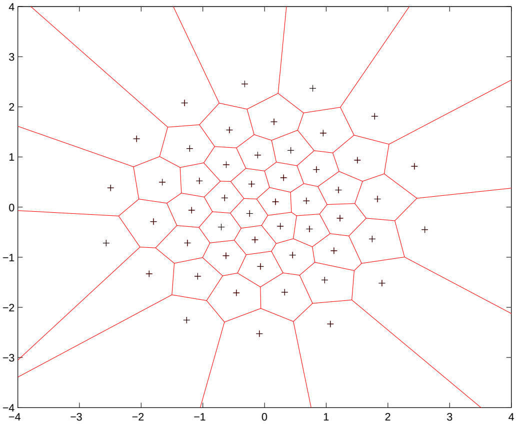


Figure: Optimal quantizer and tessellation of a 2-d Gaussian r.v.

Vector quantization: some useful facts

Theory:

- The L^r- **error** goes to zero as $N \to +\infty$ (**Zador Theorem**).
- The distortion function (the quadratic quantization error squared) always reaches one minimum at a N-tuple $\Gamma^\star$ having pairwise distinct components.

Practice:

- $d = 1$: optimal quantizers can be obtained via standard Newton-Raphson procedure.
- $d \geq 2$: stochastic gradient descent algorithms are required (or standard gradient descent when the distribution can be easily simulated)

Vector quantization: fixing the ideas

Optimal quadratic quantization of X:

gives access to a N-tuple $\Gamma = \{x_1, x_2, \ldots, x_N\}$
which minimizes the L^2 distance
between X and $\widehat{X}^\Gamma$

This provides the best possible quadratic approximation of a random vector X by a random vector taking (at most) N values.

Vector quantization: numerical integration

Given an integrable function f, a random variable X and a (hopefully optimal) quantizer $\Gamma = \{x_1, \dots, x_N\}$, $\mathbb{E}[f(X)]$ can be approximated by the finite sum

$$\mathbb{E}[f(\widehat{X}^\Gamma)] = \sum_{i=1}^{N} f(x_i)\mathbb{P}(\widehat{X}^\Gamma = x_i).$$

If f is Lipschitz continuous, then

$$\left|\mathbb{E}[f(X)] - \mathbb{E}[f(\widehat{X^\Gamma})]\right| \leq [f]_{\text{Lip}} \, ||X - \widehat{X}^\Gamma||_2$$

and $||X - \widehat{X}^\Gamma||_2 \overset{N\to\infty}{\longrightarrow} 0$ (Zador theorem).

N.B. When f is smoother this error bound can be significantly improved.

Vector quantization: towards stationary quantizers

What do we need in practice to quantize X?

- The grid $\Gamma^\star = \{x_1, x_2, \dots, x_N\}$
- The weights of the cells in the Voronöi tessellation $\mathbb{P}(X \in C_i(\Gamma^\star)) = \mathbb{P}(\widehat{X} = x_i), i = 1, \dots, N$

From a numerical point of view, finding an optimal quantizer may be a very challenging and time consuming task.
This motivates the introduction of sub-optimal criteria: stationary quantizers.

Stationary quantizers

- **Definition:** $\Gamma = \{x_1, \cdots, x_N\}$ is stationary for X if

$$\mathbb{E}\left[X|\widehat{X}^{\Gamma}\right] = \widehat{X}^{\Gamma}.$$

- Optimal quantizers are stationary;
- Stationary quantizers $\overline{\Gamma}$ are critical points of the distortion function:

$$\boxed{\nabla D(\overline{\Gamma}) = 0} \tag{1}$$

where the distortion function is the square of the L^2-error

$$D(\Gamma) := \sum_{i=1}^{N} \int_{C_i(\Gamma)} |u - x_i|^2 d\mathbb{P}_X(u).$$

From stationary quantizers to the quantization of a stochastic process

Stationary quantizers are interesting from a numerical point of view: they can be found through zero search recursive procedures like Newton's algorithm.

$$\Downarrow$$

- We ♡ stationary (sub-optimal) quantizers.
- How to quantize a stochastic process with these ideas?

"Step by step marginal quantization": warm up

Recently introduced by Pagès and Sagna [8]

- Consider a continuous-time Markov process Y

$$dY_t = b(t, Y_t)dt + a(t, Y_t)dW_t, \quad Y_0 = y_0 > 0,$$

  where W is a standard Brownian motion and a and b satisfy the usual conditions ensuring the existence of a (strong) solution to the SDE;

- Given $T > 0$ and $\{0 = t_0, t_1, \ldots, t_M = T\}$, $\Delta_k = t_k - t_{k-1}$, $k \geq 1$, the **Euler scheme** is

$$\widetilde{Y}_{t_k} = \widetilde{Y}_{t_{k-1}} + b(t_{k-1}, \widetilde{Y}_{t_{k-1}})\Delta_k + a(t_{k-1}, \widetilde{Y}_{t_{k-1}})\Delta W_k$$
$$\widetilde{Y}_{t_0} = \widetilde{Y}_0 = y_0$$

  where $\Delta W_k := (W_{t_k} - W_{t_{k-1}}) \sim \mathcal{N}(0, \Delta_k)$.

- **Key remark:** for every $k = 1, \dots, M$

$$\boxed{\mathscr{L}\left(\widetilde{Y}_{t_k} \,\middle|\, \widetilde{Y}_{t_{k-1}} = x\right) \sim \mathscr{N}\left(m_{k-1}(x), \sigma^2_{k-1}(x)\right)} \tag{2}$$

  where

$$\begin{aligned} m_{k-1}(x) &= x + b(t_{k-1}, x)\Delta_k \\ \sigma^2_{k-1}(x) &= \left[a(t_{k-1}, x)\right]^2 \Delta_k. \end{aligned}$$

- **Idea:** quantize recursively every marginal random variable (vector quantization) $\widetilde{Y}_{t_k}$, exploiting (2).

"Step by step marginal quantization": stationary quantizers

The distortion function at time t_k, relative to $\widetilde{Y}_{t_k}$, is

$$D_k(\Gamma^k) = \sum_{i=1}^{N} \int_{C_i(\Gamma^k)} (y - y_i^k)^2 \; \mathbb{P}\big(\widetilde{Y}_{t_k} \in dy\big)$$

where N is the (fixed) size of the grid $\Gamma^k = \{y_1^k, y_2^k, \ldots, y_N^k\}$. Target: $\Gamma^k \in \mathbb{R}^N$ such that

$$\boxed{\nabla D_k(\Gamma^k) = 0}$$

Question: applying Newton-Raphson now?
Answer: NO! We do NOT know the distribution of $\widetilde{Y}_{t_k}$!

- ... using the conditional distribution in (2) we have

$$\mathbb{P}(\widetilde{Y}_{t_k} \in dy) = dy \int_{\mathbb{R}} \phi_{m_{k-1}(y_{k-1}),\sigma_{k-1}(y_{k-1})}(y)\ \mathbb{P}(\widetilde{Y}_{t_{k-1}} \in dy_{k-1})$$

where $\phi_{m,\sigma}$ is the density function of a $\mathcal{N}(m,\sigma^2)$.
- Replacing $\widetilde{Y}$ by $\widehat{Y}$, we deduce a recursive procedure to obtain the stationary quantizer at time t_k, based on the quantizer at time $t_{k-1}, k \in \{0,\ldots,M-1\}$:

The distorsion is continuously differentiable, so (via gradient and Hessian matrix) $\rightsquigarrow$ Newton-Raphson $\rightsquigarrow$ faster computations wrt stochastic algorithms.

The algorithm

At every step $k = 1, \dots, M-1$ of the algorithm:

- **What we need**
  - The (stationary) quantizer $\widehat{Y}_{k-1}$ at time t_{k-1}.
  - The weights
- **What we do**

  Newton - Raphson iterations until convergence to the stationary grid $\Gamma^k = (y_1^k, \dots, y_N^k)$ at time t_k .
- **What we get**
  - The quantization at time t_k:

$$\widehat{Y}_k = \sum_{i=1}^{N} y_i^k \mathbb{1}_{\widetilde{Y}_k \in C_i(\Gamma^k)}.$$

  - The weights
  - The transition probabilities from time t_{k-1} to time t_k.

Recent research and perspectives - 1

Recursive marginal quantization can be safely extended to discretize Y taking values in $\mathbb{R}^d \rightsquigarrow$ (local and) stochastic vola models ($d = 2$).
Example: Heston model

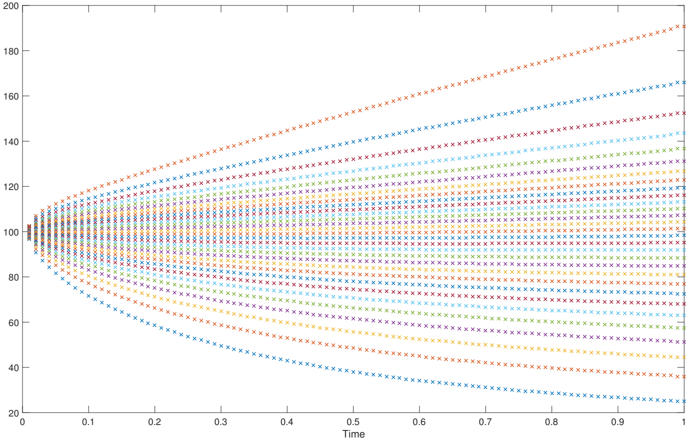
$$\frac{dS_t}{S_t} = rdt + \sqrt{V_t}(\rho \, dW_t^1 + \sqrt{1-\rho^2} dW_t^2)$$

$$dV_t = \kappa(\theta - V_t)dt + \xi\sqrt{V_t} dW_t^1$$

where

- W^1 and W^2 are independent standard Brownian motions
- r is the risk free interest rate
- θ is the long run average price variance
- κ is the reversion speed
- ρ is the correlation
- ξ is the volatility of the variance process (vol of vol).

RMQuantization of the Heston model



Quantization of the price process in the Heston model, $N = 30$.

Recent research and perspectives - 2

- Being transition probabilities available, it is also possible to easily price exotic options (such as American).
- Calibration on vanilla and american options' prices is possible.
- Challenges:
  - High dimension $\rightsquigarrow$ machine learning?
  - Discretizing non Markovian stochastic processes (e.g. rough-volatility models)?

- Bormetti, G., Callegaro, G., Livieri, G. and Pallavicini, A. (2018). A backward Monte Carlo approach to exotic option pricing, European Journal of Applied Mathematics, 29(1), pp. 146-187.
- Callegaro, G., Fiorin, L. and Grasselli, M. (2015). Quantized calibration in local volatility. Risk, 56-67.
- Callegaro, G., Fiorin, L. and Grasselli, M. (2016). Pricing via quantization in stochastic volatility models, Quantitative Finance, 17(6), pp. 855-872.
- Callegaro, G., Fiorin, L. and Grasselli, M. (2018). American quantized calibration in stochastic volatility, Risk, February 2018, pp. 84-88.
- Callegaro, G., Fiorin, L. and Grasselli, M. (2018). Quantization meets Fourier: a new technology for pricing options, Annals of Operations Research, to appear.
- Callegaro, G., Fiorin, L. and Pallavicini, A. (2018). Quantization goes polynomial, preprint.
- Graf, S. and Luschgy, H. (2000). Foundations of quantization for probability distributions. Springer, New York.
- Pagès, G. and Sagna, A. (2015). Recursive marginal quantization of the Euler scheme of a diffusion process. Applied Mathematical Finance, 22(5), pp. 463-498.
- Mc Walter, T.A., Rudd, R., Kienitz, J. and Platen, E. (2018). Recursive Marginal Quantization of Higher-Order Schemes. Quantitative Finance, 18(4), pp. 693-706.

Thank you for your attention !!!