Model uncertainty in finance :
Term-structure construction and hedging issues

Areski Cousin
ISFA, Université Lyon 1

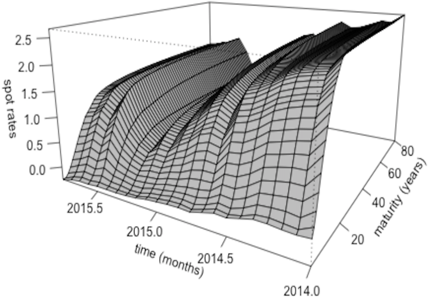
Séminaire équipe TOSCA

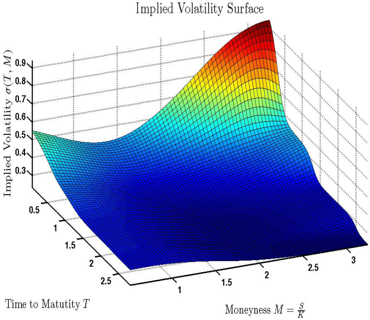
INRIA Sophia-Antipolis, November 24, 2016

Contents

Motivation

- Financial term-structures describes the evolution of some financial or economic quantities as a function of time horizon.
- **Examples** : interest-rates, bond yields, credit spreads, implied default probabilities, implied volatilities.
- **Applications** : valuation of financial and insurance products, risk management





The term-structure construction problem

Several constraints have to be considered

- Compatibility with market information : at a given date t_0, the curve under construction $T \to P(t_0, T)$ shall be compatible with observed prices of some reference products.
- Arbitrage-free construction : this translates into some specific shape properties such as positivity, monotonicity, convexity or bounds on the curve values
- Additional conditions can be required : minimum degree of smoothness, control of local convexity

The term-structure construction problem

1) Compatibility with market information :

- At time t_0, we observe the market quotes $S_1, \ldots, S_n$ of n liquidly traded instruments
- The values of these products depend on the value of the curve at m input locations $X = (\tau_1, \ldots, \tau_m)$

The vector of output values $P(t_0, X) := (P(t_0, \tau_1), \ldots, P(t_0, \tau_m))^\top$ satisfies a linear system of the form

$$A \cdot P(t_0, X) = \boldsymbol{b},$$

where

- A is a $n \times m$ real-valued matrix
- $\boldsymbol{b}$ is a n-dimensional column vector

$n < m \Longrightarrow$ indirect and partial information on the curve values at $\tau_1, \ldots, \tau_m$

2) No-arbitrage assumption :

$T \to P(t_0, T)$ is typically a monotonic bounded function

Range of arbitrage-free OIS discount curves

We observe the quoted par rates S_i of an OIS with maturities T_i, $i = 1, \ldots, n$

1) Compatibility with market quotes :

The curve $T \to P(t_0, T)$ of **OIS discount factors** is such that

$$S_i \sum_{k=1}^{p_i - 1} \delta_k P(t_0, t_k) + (S_i \delta_{p_i} + 1)\, P(t_0, T_i) = 1, \quad i = 1, ..., n$$

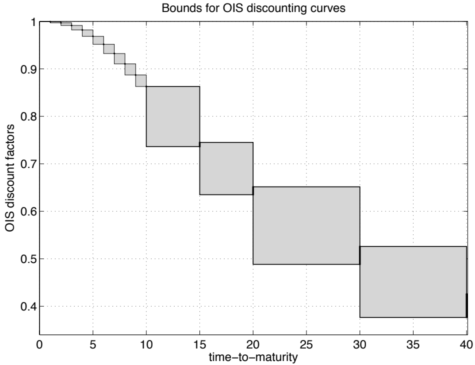
- $t_1 < \cdots < t_{p_i} = T_i$: fixed-leg payment dates (annual time grid)
- δ_k : year fraction of period (t_{k-1}, t_k)

2) No-arbitrage assumption :

$T \to P(t_0, T)$ is a decreasing function such that $P(t_0, t_0) = 1$

Range of arbitrage-free OIS discount curves

- $n = 14$ liquidly traded maturities $1, 2, \ldots, 10, 15, 20, 30, 40$ years.
- $m = 40$ points involved in the market-fit linear system
- No-arbitrage bounds on OIS discount factors



Input data : OIS swap rates as of May, 31st 2013.
Source : Cousin and Niang (2014)

Range of arbitrage-free CDS-implied survival functions

We observe at time t_0 the fair spreads S_i of a CDS with maturities T_i, $i = 1, \ldots, n$

1) Compatibility with market quotes :

The curve $T \to P(t_0, T)$ of (risk-neutral) **survival probabilities** is such that

$$S_i \sum_{k=1}^{p_i} \delta_k D(t_0, t_k) P(t_0, t_k) = -(1-R) \int_{t_0}^{T_i} D(t_0, u) dP(t_0, u), \quad i = 1, \ldots, n$$

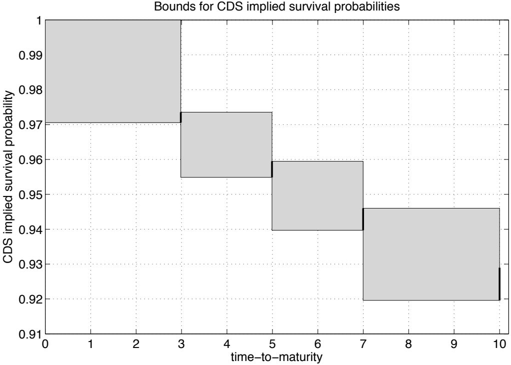
- $t_1 < \cdots < t_p = T_i$: trimestrial premium payment dates, δ_k : year fraction of period (t_{k-1}, t_k)
- $D(t_0, T)$ is the discount factor associated with maturity date T
- R : expected recovery rate of the reference entity

2) No-arbitrage assumption :

$T \to P(t_0, T)$ is a decreasing function such that $P(t_0, t_0) = 1$

Range of arbitrage-free CDS-implied survival functions

- $n = 4$ liquidly traded maturities $3, 5, 7, 10$ years.
- $m = 40$ points involved in the market-fit linear system
- No-arbitrage bounds on the issuer implied survival distribution function



Input data : CDS spreads of AIG as of December 17, 2007, $R = 40\%$,
$D(t, T) = \exp(-3\%(T - t))$

From spline interpolation to kriging

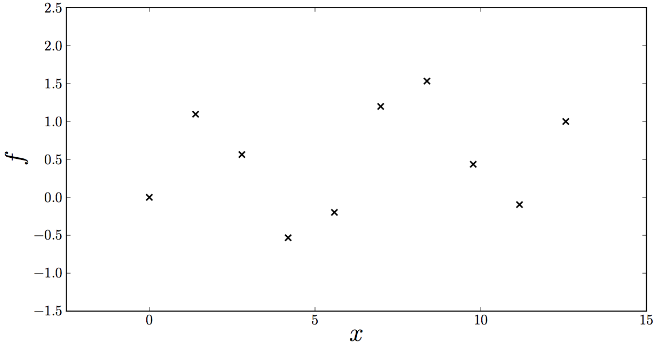
In practice, financial term-structures are constructed using deterministic interpolation techniques.

- Parametric approaches : Nelson-Siegel or Svensson models (used by most central banks)
- Non-parametric interpolation methods : shape-preserving spline techniques (lack of interpretability but better ability to fit the data).

Could we propose an arbitrage-free interpolation method that additionally allows for quantification of uncertainty ?

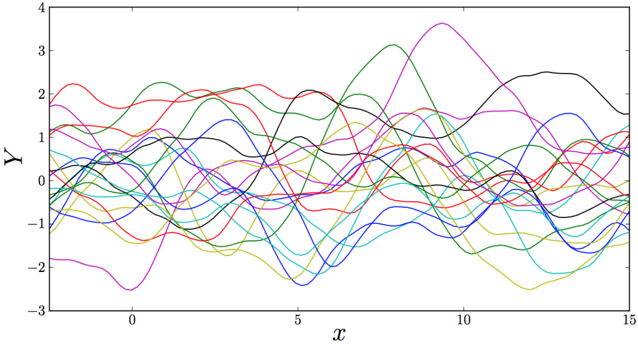
Classical kriging

A function f is only known at a limited number of points $x_1, \dots, x_n$



Classical kriging

The (unknown) function f is assumed to be a sample path of a Gaussian process Y



Classical kriging

Definition : Gaussian process (GP) or Gaussian random field

A Gaussian process is a collection of random variables, any finite number of which have (consistent) joint Gaussian distributions.

A Gaussian process $\left(Y(x), x \in \mathbb{R}^d\right)$ is characterized by its mean function

$$\mu : x \in \mathbb{R}^d \longrightarrow \mathbb{E}(Y(x)) \in \mathbb{R}.$$

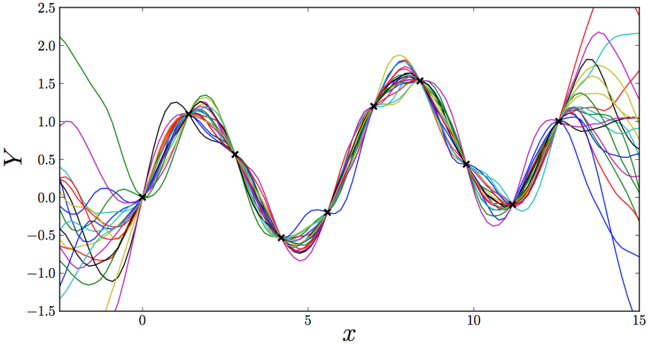
and its covariance function

$$K : (x, x') \in \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathrm{Cov}(Y(x), Y(x')) \in \mathbb{R}.$$

1D kriging kernel	$K(x, x')$	Class
Gaussian	$\sigma^2 \exp\left(-\frac{(x-x')^2}{2\theta^2}\right)$	$\mathcal{C}^\infty$
Matérn 5/2	$\sigma^2 \left(1 + \frac{\sqrt{5}\lvert x-x'\rvert}{\theta} + \frac{5(x-x')^2}{3\theta^2}\right) \exp\left(-\frac{\sqrt{5}\lvert x-x'\rvert}{\theta}\right)$	$\mathcal{C}^2$
Matérn 3/2	$\sigma^2 \left(1 + \frac{\sqrt{3}\lvert x-x'\rvert}{\theta}\right) \exp\left(-\frac{\sqrt{3}\lvert x-x'\rvert}{\theta}\right)$	$\mathcal{C}^1$
Exponential	$\sigma^2 \exp\left(-\frac{\lvert x-x'\rvert}{\theta}\right)$	$\mathcal{C}^0$

Classical kriging

The estimation of f relies on the conditional distribution of Y given the observed values $y_i = f(x_i)$ at points x_i, $i = 1, \ldots, n$.



Classical kriging

- $\boldsymbol{X} = (x_1, \ldots, x_n)^\top \in \mathbb{R}^{n\times d}$: some design points
- $\boldsymbol{y} = (y_1, \ldots, y_n)^\top \in \mathbb{R}^n$: observed values of f at these points
- $Y(\boldsymbol{X}) = (Y(x_1), \ldots, Y(x_n))^\top$: vector composed of Y at point $\boldsymbol{X}$

The conditional process is still a Gaussian Process

Let Y be a GP with mean μ and covariance function K. The conditional process $Y \mid Y(\boldsymbol{X}) = \boldsymbol{y}$ is a GP with mean function

$$\eta(x) = \mu(x) + \boldsymbol{k}(x)^\top \mathbb{K}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}), \quad x \in \mathbb{R}^d$$

and covariance function $\tilde{K}$ given by

$$\tilde{K}(x, x') = K(x, x') - \boldsymbol{k}(x)^\top \mathbb{K}^{-1}\boldsymbol{k}(x'), \quad x, x' \in \mathbb{R}^d$$

where $\boldsymbol{\mu} = \mu(X) = (\mu(x_1), \ldots, \mu(x_n))^\top$, $\mathbb{K}$ is the covariance matrix of $Y(\boldsymbol{X})$ and $\boldsymbol{k}(x) = (K(x, x_1), \ldots, K(x, x_n))^\top$

Extension to linear equality constraints

Recall that, in our term-structure construction problem, the (unknown) real function f satisfies some linear equality constraints of the form

$$A \cdot f(X) = \boldsymbol{b}, \tag{1}$$

where

- A is a given matrix of dimension $n \times m$
- $X = (x_1, \ldots, x_m)^\top \in \mathbb{R}^{m \times d}$
- $f(X) = (f(x_1), \ldots, f(x_m))^\top \in \mathbb{R}^m$
- $\boldsymbol{b} \in \mathbb{R}^n$

Extension to linear equality constraints

- $X = (x_1, \ldots, x_m)^\top \in \mathbb{R}^{m \times d}$: some design points
- $\boldsymbol{b} = (b_1, \ldots, b_n)^\top \in \mathbb{R}^n$: right-hand side of the linear system
- $Y(X) = (Y(x_1), \ldots, Y(x_m))$: vector composed of Y at point $\boldsymbol{X}$

The conditional process is still a Gaussian Process

Let Y be a GP with mean μ and covariance function K. The conditional process $Y \mid AY(X) = \boldsymbol{b}$ is a GP with mean function

$$\eta(x) = \mu(x) + (A\boldsymbol{k}(x))^\top \left(A\mathbb{K}A^\top\right)^{-1} (\boldsymbol{b} - A\boldsymbol{\mu}), \quad x \in \mathbb{R}^d$$

and covariance function $\tilde{K}$ given by

$$\tilde{K}(x, x') = K(x, x') - (A\boldsymbol{k}(x))^\top \left(A\mathbb{K}A^\top\right)^{-1} A\boldsymbol{k}(x'), \quad x, x' \in \mathbb{R}^d$$

where $\boldsymbol{\mu} = \mu(X) = (\mu(x_1), \ldots, \mu(x_m))^\top$, $\mathbb{K}$ is the covariance matrix of $Y(X)$, $\boldsymbol{k}(x) = (K(x, x_1), \ldots, K(x, x_m))^\top$

Extension to monotonicity constraints

New formulation of the problem : estimation of an unknown function f given that

$$\begin{cases} A \cdot f(X) = \boldsymbol{b} \\ f \in \mathcal{M} \end{cases}$$

where $\mathcal{M}$ is the set of (say) non-increasing functions.

Problem : The conditional process is not a Gaussian process anymore.

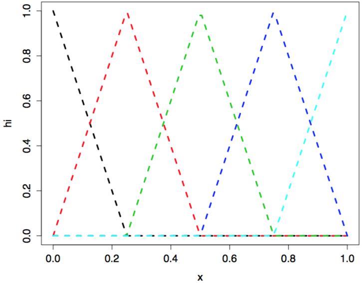
- How to cope with the infinite-dimensional monotonicity constraints ?
- Which estimator could we propose for the term-structure ?

Extension to monotonicity constraints (1D case)

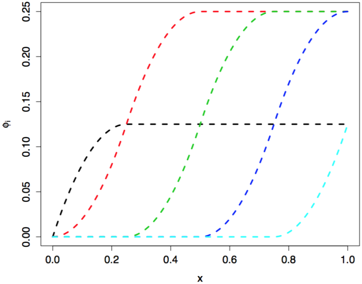
Proposed methodology : On an interval $D = [\underline{x}, \overline{x}]$ of $\mathbb{R}$, we construct a finite-dimensional approximation of Y for which the monotonicity constraint is easy to check.

- Regular subdivision $u_0 < \ldots < u_N$ of D with a constant mesh δ
- Set of increasing basis functions $(\phi_i)_{i=0,\ldots,N}$ defined on this subdivision

$$h_i(x) := \max\left(1 - \frac{|x-u_i|}{\delta}, 0\right)$$



$$\phi_i(x) = \int_{\underline{x}}^{x} h_i(u)du$$



Extension to monotonicity constraints (1D case)

Proposition (Maatouk and Bay, 2014b)

Let Y be a zero-mean GP with covariance function K and with almost surely differentiable paths.

- The finite-dimensional process Y^N defined on D by

$$Y^N(x) = Y(u_0) + \sum_{j=0}^{N} Y'(u_j)\phi_j(x)$$

  uniformly converges to Y, almost surely.

- Y^N is non-decreasing (resp. non-increasing) on D if and only if $Y'(u_j) \geq 0$ (resp. $Y'(u_j) \leq 0$) for all $j = 0, \dots, N$.

- Let $\xi := (Y(u_0), Y'(u_0), \dots, Y'(u_N))^\top$, then $\xi \sim \mathcal{N}(0, \Gamma^N)$ where

$$\Gamma^N = \begin{bmatrix} K(u_0, u_0) & \frac{\partial K}{\partial x'}(u_0, u_j) \\ \frac{\partial K}{\partial x}(u_i, u_0) & \frac{\partial^2 K}{\partial x \partial x'}(u_i, u_j) \end{bmatrix}_{0 \leq i,j \leq N}$$

Extension to monotonicity constraints (1D case)

For a given covariance function K, we assume that the unknown function f is a sample path of the GP

$$Y^N(x) = \eta + \sum_{j=0}^{N} \xi_j \phi_j(x), \qquad x \in D,$$

where $\boldsymbol{\xi} := (\eta, \xi_0, \ldots, \xi_N)^\top \sim \mathcal{N}(0, \Gamma^N)$.

Kriging f is equivalent to find the conditional distribution of Y^N given

$$\begin{cases} A \cdot Y^N(X) = \boldsymbol{b} & \text{linear equality condition} \\ \xi_j \leq 0,\ j = 0, \ldots, N & \text{monotonicity constraint} \end{cases}$$

Extension to monotonicity constraints (1D case)

Or equivalently, to find the distribution of the truncated Gaussian vector $\boldsymbol{\xi} \sim \mathcal{N}(0, \Gamma^N)$ given

$$\begin{cases} A \cdot \Phi \cdot \boldsymbol{\xi} = \boldsymbol{b} & \text{linear equality condition} \\ \xi_j \leq 0, \; j = 0, \ldots, N & \text{monotonicity constraint} \end{cases}$$

where Φ is a $m \times (N+2)$ matrix defined as

$$\Phi_{i,j} := \begin{cases} 1 & \text{for } i = 1, \ldots, m \text{ and } j = 1, \\ \phi_{j-2}(x_i) & \text{for } i = 1, \ldots, m \text{ and } j = 2, \ldots, N+2. \end{cases}$$

Extension to monotonicity constraints (1D case)

Which estimator could we use for f ?

We consider the mode of the truncated gaussian process (most probable path) :

$$M_K^N(x \mid A, \boldsymbol{b}) = \nu + \sum_{j=0}^{N} \nu_j \phi_j(x),$$

where $\boldsymbol{\nu} = (\nu, \nu_0, \ldots, \nu_N)^\top \in \mathbb{R}^{N+2}$ is the solution of the following convex optimization problem :

$$\boldsymbol{\nu} = \arg \min_{\boldsymbol{c} \in \mathcal{C} \cap \mathcal{I}(A,\boldsymbol{b})} \left(\frac{1}{2} \boldsymbol{c}^\top \left(\Gamma^N\right)^{-1} \boldsymbol{c} \right),$$

with

- $\mathcal{C} = \left\{ \boldsymbol{\xi} \in \mathbb{R}^{N+2} \ : \ \xi_j \leq 0, \ j = 0, \ldots, N \right\}$
- $\mathcal{I}(A, \boldsymbol{b}) = \left\{ \boldsymbol{\xi} \in \mathbb{R}^{N+2} \ : \ A \cdot \Phi \cdot \boldsymbol{\xi} = \boldsymbol{b} \right\}$

Extension to monotonicity constraints (1D case)

Efficient simulation of the truncated Gaussian vector

1) Simulate a truncated vector $\boldsymbol{\xi}$ given the linear equality constraint :

$$Z \sim \{\boldsymbol{\xi} \mid B \cdot \boldsymbol{\xi} = \boldsymbol{b}\} \sim \mathcal{N}\left((B\Gamma^N)^\top \left(B\Gamma^N B^\top\right)^{-1} \boldsymbol{b}, \Gamma^N - \left(B\Gamma^N\right)^\top \left(B\Gamma^N B^\top\right)^{-1} B\Gamma^N\right)$$

where $B = A \cdot \Phi$.

2) Simulate

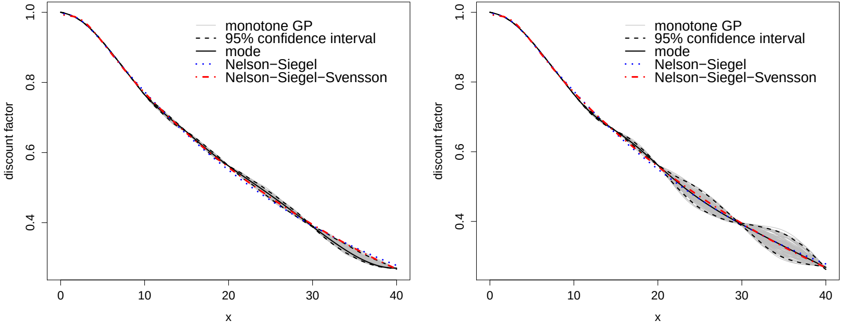
$$\{Z \mid \xi_j \leq 0, j = 0, \ldots, N\} \sim \{\boldsymbol{\xi} \mid B \cdot \boldsymbol{\xi} = \boldsymbol{b} \text{ and } \xi_j \leq 0, j = 0, \ldots, N\}$$

by an accelerated rejection sampling method (we use the method proposed in Maatouk and Bay, 2014a)

3) The corresponding sample curves $Y^N(\cdot) = \eta + \sum_{j=0}^{N} \xi_j \phi_j(\cdot)$ satisfies the constraints on the entire domain D.

Kriging of OIS discount curves

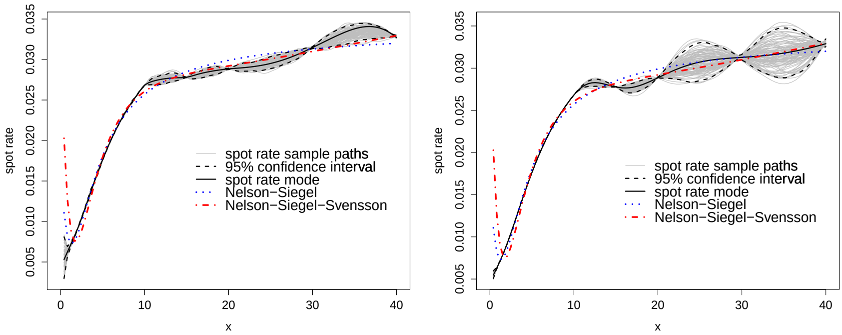
- We compare two covariance functions : Gaussian and Matérn 5/2
- Hyper-parameters θ and σ are estimated using cross-validation
- Comparison with Nelson-Siegel and Svensson curve fitting



Discount curves. $N = 50$, 100 sample paths. Left : Gaussian covariance function. Right : Matérn 5/2 covariance function. OIS data of 03/06/2010.

Kriging of OIS discount curves

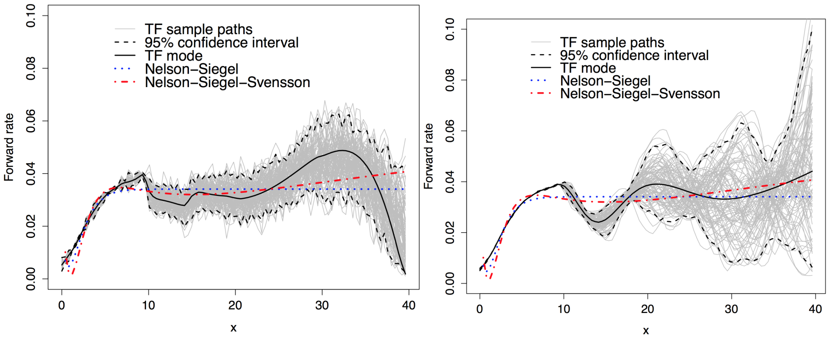
Corresponding spot rate curves : $-\frac{1}{x} \log P(x)$



Spot rate curves. Left : Gaussian covariance function. Right : Matérn 5/2 covariance function. OIS data of 03/06/2010. The black solid line is the most likely spot rate curve $-\frac{1}{x} \log M_K^N(x \mid A, \boldsymbol{b})$.

Kriging of OIS discount curves

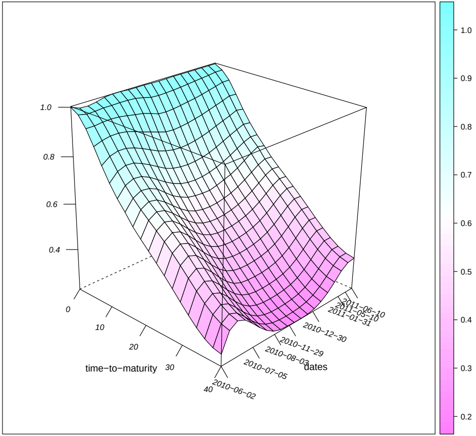
Corresponding forward rate curves : $-\frac{d}{dx}\log P(x)$



Spot rate curves. Left : Gaussian covariance function. Right : Matérn 5/2 covariance function. OIS data of 03/06/2010. The black solid line is the most likely forward rate curve $-\frac{d}{dx}\log M_K^N(x \mid A, \boldsymbol{b})$.

Kriging of OIS discount curves (2D)

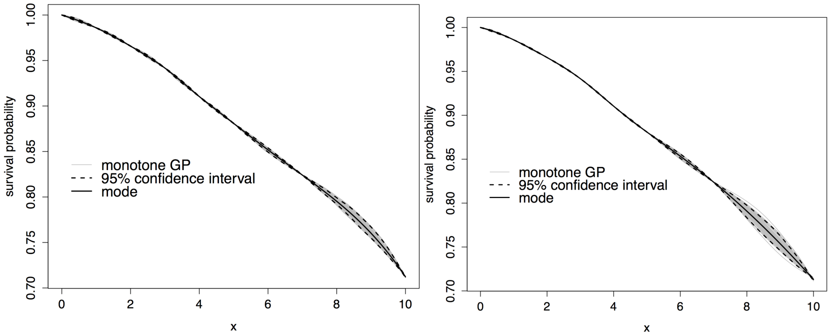
The previous approach can be extended in dimension 2.



Dicount curves. OIS discount factors as a function of time-to-maturities and quotation dates.

Kriging of CDS-implied default distribution

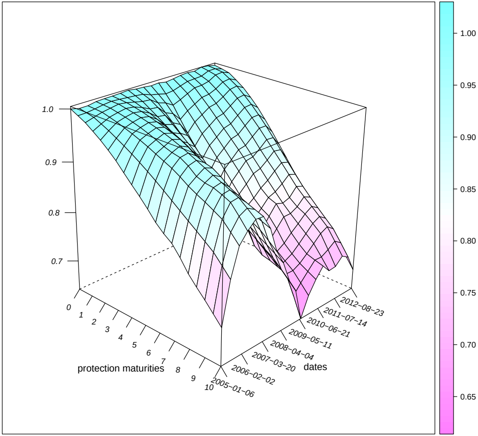
Implied survival function of the Russian sovereign debt



CDS implied survival curves. $N = 50, 100$ sample paths. Left : Gaussian covariance function. Right : Matérn 5/2 covariance function. CDS spreads as of 06/01/2005.

Kriging of CDS-implied default distribution (2D)

The previous approach can be extended in dimension 2.

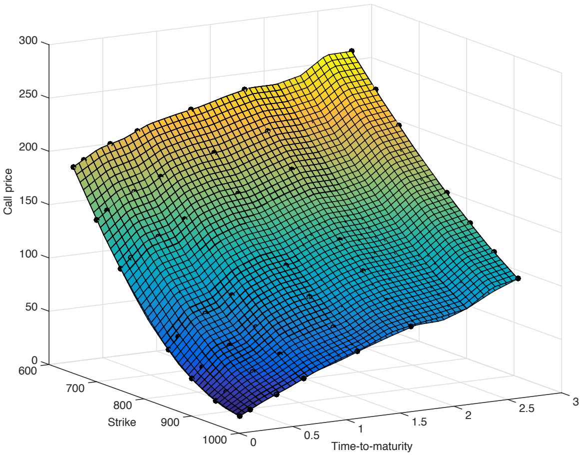


Survival curves. CDS implied survival probabilities as a function of time-to-maturities and quotation dates.

Perspectives

- Impact of curve uncertainty on the assessment of related products and their associated hedging strategies
- What if the underlying market quotes are not reliable due to e.g. market illiquidity (data observed with a noise) ?
- Kriging of arbitrage-free volatility surfaces ?

Kriging of arbitrage-free volatility surface



Contents

General setting and motivation

- Robust control may be overly conservative when applied to the true unknown system
- We develop an adaptive robust methodology for solving a discrete-time Markovian control problem subject to Knightian uncertainty
- We focus on a financial hedging problem, but the methodology can be applied to any kind of Markov decision process under model uncertainty
- As in the classical robust case, the uncertainty comes from the fact that the true law of the driving process is only known to belong to a certain family of probability laws

General setting and motivation

- T : terminal date of our finite horizon control problem
- $\mathcal{T} = \{0, 1, 2, \dots, T\}$: time grid
- $\mathcal{T}' = \{0, 1, 2, \dots, T-1\}$: time grid without last date
- $S = \{S_t,\ t \in \mathcal{T}\}$: stochastic process that drives the random system

We assume that :

- S is observable and we denote by $\mathbb{F}^S = (\mathscr{F}_t^S, t \in \mathcal{T})$ its natural filtration.
- The law of S is not known but it belongs to a family of parametrized distributions $\mathbf{P}(\mathbf{\Theta}) := \{\mathbb{P}_\theta, \theta \in \mathbf{\Theta}\}$, $\mathbf{\Theta} \subset \mathbb{R}^d$
- The unknown (true) law of S is denoted by $\mathbb{P}_{\theta^*}$ and is such that $\theta^* \in \mathbf{\Theta}$

Model uncertainty occurs if $\mathbf{\Theta} \neq \{\theta^*\}$

General setting and motivation

We consider the following stochastic control problem

$$\inf_{\varphi \in \mathcal{A}} \mathbb{E}_{\theta^*}\left(L(S,\varphi)\right).$$

where

- $\mathcal{A}$ is a set of admissible control processes : $\mathbb{F}^S$–adapted processes $\varphi = \{\varphi_t,\ t \in \mathcal{T}'\}$
- L is a measurable functional (loss or error to minimize in our case)

Obviously, the problem cannot be dealt with directly since we do not know the value of θ^*

General setting and motivation

Robust control problem : Başar and Bernhard (1995), Hansen et al. (2006), Hansen and Sargent (2008)

$$\inf_{\varphi \in \mathcal{A}} \sup_{\theta \in \boldsymbol{\Theta}} \mathbb{E}_{\theta} \left(L(S, \varphi) \right). \tag{2}$$

- Best strategy over the worst possible model parameter in $\boldsymbol{\Theta}$
- If the true model is close to the best one, the solution to this problem could perform very badly

General setting and motivation

Strong robust control problem : Sirbu (2014), Bayraktar, Cosso and Pham (2014)

$$\inf_{\varphi \in \mathcal{A}} \sup_{\mathbb{Q} \in \mathcal{Q}^{\varphi, \boldsymbol{\Psi}_K}} \mathbb{E}_{\mathbb{Q}} \left(L(S, \varphi) \right), \tag{3}$$

- $\boldsymbol{\Psi}_K$ is the set of strategies chosen by a Knightian adversary (the nature) that may keep changing the system distribution over time
- $\mathcal{Q}^{\varphi, \boldsymbol{\Psi}_K}$ represents all possible model dynamics resulting from φ and when nature plays strategies in $\boldsymbol{\Psi}_K$
- Solution is even more conservative than in the classical robust case
- No learning mechanism to reduce model uncertainty

General setting and motivation

Adaptive control problem : Kumar and Varaiya (1986), Chen and Guo (1991)

For each $\theta \in \mathbf{\Theta}$ solve :

$$\inf_{\varphi \in \mathcal{A}} \mathbb{E}_{\theta} \left(L(S, \varphi) \right). \tag{4}$$

- Let φ^{θ} be a corresponding optimal control
- At each time t, compute a point estimate $\widehat{\theta}_t$ of θ^*, using a chosen, $\mathscr{F}_t^S$ measurable estimator and apply control value $\varphi_t^{\widehat{\theta}_t}$.
- Known to have poor performance for finite horizon problems

Hedging under model uncertainty

Problem : Hedging a short position on an European-type option with maturity T, payoff function Φ and underlying asset S with price dynamics

$$
\begin{aligned}
S_0 &= s_0 \in (0, \infty), \\
S_{t+1} &= Z_{t+1} S_t, \qquad t \in \mathcal{T}'
\end{aligned}
$$

where

- $Z = \{Z_t,\ t = 1, \dots, T\}$ is a non-negative random process
- Under each measure $\mathbb{P}_\theta$, Z_{t+1} is independent from $\mathscr{F}_t^S$ for each $t \in \mathcal{T}$
- The true law $\mathbb{P}_{\theta^*}$ of Z is not known.

Hedging under model uncertainty

Hedging is made using a self-financing portfolio composed of the underlying risky asset S and of a risk-free asset (with constant value equal to 1).

The hedging portfolio has the following dynamics

$$
\begin{aligned}
V_0 &= v_0, \\
V_{t+1} &= V_t + \varphi_t(S_{t+1} - S_t), \quad t = 0, \ldots, T-1
\end{aligned}
$$

Exact replication is out of reach in our setting (v_0 may be too small), so that the nominal control problem (without uncertainty) is

$$\inf_{\varphi \in \mathcal{A}} \mathbb{E}_{\theta^*} \left(\ell[(\Phi(S_T) - V_T(\varphi))^+] \right),$$

where l is a loss function, i.e., an increasing function such that $\ell(0) = 0$ (shortfall risk minimization approach)

Adaptive robust control methodology

The methodology relies on recursive construction of confidence regions. We assume that :

1) A point estimator $\widehat{\theta}_t$ of θ^* can be constructed recursively

$$\widehat{\theta}_0 = \theta_0,$$
$$\widehat{\theta}_{t+1} = R(t, \widehat{\theta}_t, Z_{t+1}), \quad t = 0, \dots, T-1$$

where $R(t, c, z)$ is a deterministic measurable function.

2) An approximate α-confidence region $\mathbf{\Theta}_t$ of θ^* can be constructed from $\widehat{\theta}_t$ by a deterministic rule :

$$\mathbf{\Theta}_t = \tau(t, \widehat{\theta}_t)$$

where $\tau(t, \cdot) : \mathbb{R}^d \to 2^{\mathbf{\Theta}}$ is a deterministic set valued function. The region $\mathbf{\Theta}_t$ should be such that $\mathbb{P}_{\theta^*}(\theta^* \in \mathbf{\Theta}_t) \approx 1 - \alpha$ and $\lim_{t\to\infty} \mathbf{\Theta}_t = \{\theta^*\}$ where the convergence is understood $\mathbb{P}^{\theta^*}$ almost surely, and the limit is in the Hausdorff metric.

Adaptive robust control methodology

We consider the following (augmented) state process

$$X_t = (S_t, V_t, \widehat{\theta}_t), \quad t \in \mathcal{T}$$

with state space $E_X := \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}^d$.

In our hedging problem, $X = (S, V, \widehat{\theta})$ is a Markov process with dynamics :

$$\begin{aligned} S_{t+1} &= Z_{t+1} S_t, \\ V_{t+1} &= V_t + \varphi_t S_t (Z_{t+1} - 1), \\ \widehat{\theta}_{t+1} &= R(t, \widehat{\theta}_t, Z_{t+1}) \end{aligned}$$

We denote by

$$Q(B \mid t, x, a, \theta) := \mathbb{P}_\theta(X_{t+1} \in B \mid X_t = x, \varphi_t = a)$$

the time-t Markov transition kernel under probability $\mathbb{P}_\theta$ when strategy a is applied

Adaptive robust control methodology

Let us denote by

$$H_t := ((S_0, V_0, \widehat{\theta}_0), (S_1, V_1, \widehat{\theta}_1), \ldots, (S_t, V_t, \widehat{\theta}_t)), \;\; t \in \mathcal{T},$$

the history of the state process up to time t.

Note that, for any admissible trading strategy φ, H_t is $\mathscr{F}_t^S$ measurable and

$$H_t \in \mathbf{H}_t := \underbrace{E_X \times E_X \times \ldots \times E_X}_{t+1 \text{ times}}.$$

We denote by

$$h_t = (x_0, x_1, \ldots, x_t) = (s_0, v_0, c_0, s_1, v_1, c_1, \ldots, s_t, v_t, c_t)$$

a realization of H_t.

Adaptive robust control methodology

A robust control problem can be viewed as a game between a controller and nature (the Knightian opponent).

The controller plays history-dependent strategies φ that belong to

$$\mathcal{A} = \{(\varphi_t)_{t \in \mathcal{T}'} \mid \varphi_t : \mathbf{H}_t \to A,\ t \in \mathcal{T}'\}$$

where φ_t is a measurable mapping.

Strong robust case : nature plays history-dependent strategies ψ that belong to

$$\mathbf{\Psi}_K = \{(\psi_t)_{t \in \mathcal{T}'} \mid \psi_t : \mathbf{H}_t \to \mathbf{\Theta},\ t \in \mathcal{T}'\}$$

Adaptive robust case : nature plays history-dependent strategies ψ that belong

to

$$\mathbf{\Psi}_A = \{(\psi_t)_{t \in \mathcal{T}'} \mid \psi_t : \mathbf{H}_t \to \mathbf{\Theta}_t,\ t \in \mathcal{T}'\}$$

where $\mathbf{\Theta}_t = \tau(t, \widehat{\theta}_t)$ is the α-confidence region of θ^* at time t

Adaptive robust control methodology

Given that the controller plays φ and nature plays ψ, using Ionescu-Tulcea theorem, we define the canonical law of the state process X on E_X^T as

$$\mathbb{Q}_{h_0}^{\varphi,\psi}(B_1,\ldots,B_T) =$$
$$\int_{B_1}\cdots\int_{B_T} Q(dx_T \mid T-1, x_{T-1}, \varphi_{T-1}(h_{T-1}), \psi_{T-1}(h_{T-1}))$$
$$\cdots Q(dx_2 \mid 1, x_1, \varphi_1(h_1), \psi_1(h_1))\, Q(dx_1 \mid 0, x_0, \varphi_0(h_0), \psi_0(h_0)).$$

For a given strategy φ, we define

$$\mathcal{Q}_{h_0}^{\varphi,\mathbf{\Psi}_K} := \{\mathbb{Q}_{h_0}^{\varphi,\psi},\ \psi \in \mathbf{\Psi}_K\}$$

and

$$\mathcal{Q}_{h_0}^{\varphi,\mathbf{\Psi}_A} := \{\mathbb{Q}_{h_0}^{\varphi,\psi},\ \psi \in \mathbf{\Psi}_A\}$$

Adaptive robust control methodology

The strong robust hedging problem :

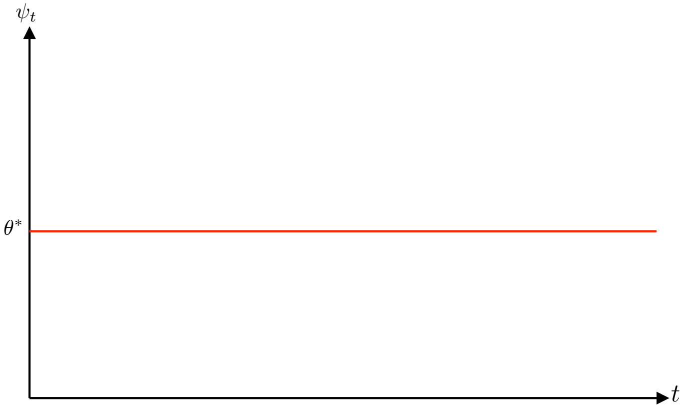
$$\inf_{\varphi \in \mathcal{A}} \sup_{\mathbb{Q} \in \mathcal{Q}_{h_0}^{\varphi, \boldsymbol{\Psi}_K}} \mathbb{E}_{\mathbb{Q}} \left(\ell[(\Phi(S_T) - V_T)^+] \right)$$

The adaptive robust hedging problem :

$$\inf_{\varphi \in \mathcal{A}} \sup_{\mathbb{Q} \in \mathcal{Q}_{h_0}^{\varphi, \boldsymbol{\Psi}_A}} \mathbb{E}_{\mathbb{Q}} \left(\ell[(\Phi(S_T) - V_T)^+] \right)$$

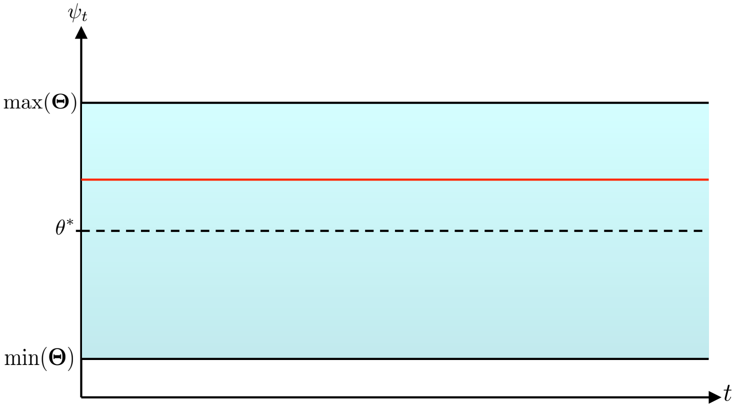
Adaptive robust control methodology

Without uncertainty



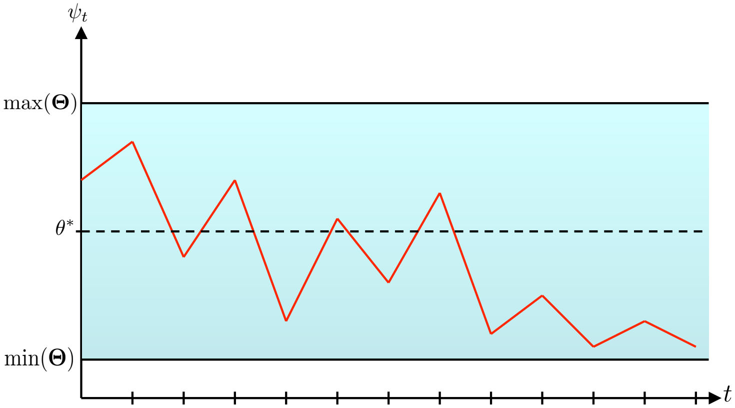
Adaptive robust control methodology

Robust



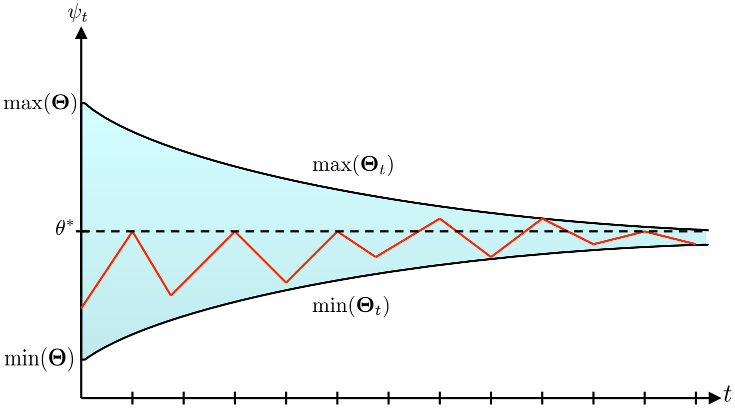
Adaptive robust control methodology

Strong robust



Adaptive robust control methodology

Adaptive robust



Adaptive robust control methodology

Dynamic programming principle

Proposition

The solution $\varphi^* = (\varphi^*_t(h_t))_{t\in\mathcal{T}'}$ of

$$\inf_{\varphi\in\mathcal{A}} \sup_{\mathbb{Q}\in\mathcal{Q}^{\varphi,\Psi_A}_{h_0}} \mathbb{E}_{\mathbb{Q}}\left(\ell[(\Phi(S_T)-V_T)^+]\right)$$

coincides with the solution of the following robust Bellman equation :

$$W_T(x) = \ell\left[(\Phi(s)-v)^+\right], \quad x=(s,v,\widehat{\theta})\in E_X,$$

$$W_t(x) = \inf_{a\in A} \sup_{\theta\in\tau(t,\widehat{\theta})} \int_{E_X} W_{t+1}(y)Q(dy \mid t,x,a,\theta),$$

for any $x=(s,v,\widehat{\theta})\in E_X$ and $t=T-1,\dots,0$.

Note that the optimal strategy at time t is such that $\varphi^*_t(h_t)=\varphi^*_t(x_t)$.

Example : uncertain log-normal model

We consider that the stock price is driven by an uncertain log-normal model

$$S_{t+1} = Z_{t+1} S_t$$

where Z_t is an iid sequence such that $\ln Z_t \overset{\mathbb{P}_{\theta^*}}{\sim} N(\mu^*, (\sigma^*)^2)$.

The MLE $\widehat{\theta}_t = (\widehat{\mu}_t, \widehat{\sigma}_t^2)$ of the unknown parameter $\theta^* = (\mu^*, (\sigma^*)^2)$ can be expressed in the following recursive way :

$$\begin{aligned}
\widehat{\mu}_{t+1} &= \frac{t}{t+1}\widehat{\mu}_t + \frac{1}{t+1}\ln Z_{t+1}, \\
\widehat{\sigma}_{t+1}^2 &= \frac{t}{t+1}\widehat{\sigma}_t^2 + \frac{t}{(t+1)^2}(\widehat{\mu}_t - \ln Z_{t+1})^2,
\end{aligned}$$

with $\widehat{\mu}_1 = \ln Z_1 = \ln \frac{S_1}{S_0}$ and $\widehat{\sigma}_1^2 = 0$.

Example : uncertain log-normal model

Due to asymptotic normality of the MLE $\widehat{\theta}_t = (\widehat{\mu}_t, \widehat{\sigma}_t^2)$, we have

$$\frac{t}{\widehat{\sigma}_t^2}(\widehat{\mu}_t - \mu^*)^2 + \frac{t}{2\widehat{\sigma}_t^4}(\widehat{\sigma}_t^2 - (\sigma^*)^2)^2 \xrightarrow[t\to\infty]{d} \chi_2^2$$

So that, if κ_α is the $(1-\alpha)$–quantile of the χ_2^2 distribution,

$$\boldsymbol{\Theta}_t = \tau(t, \widehat{\mu}, \widehat{\sigma}^2) := \left\{(\mu, \sigma^2) \in \mathbb{R}^2 : \frac{t}{\widehat{\sigma}^2}(\widehat{\mu} - \mu)^2 + \frac{t}{2\widehat{\sigma}^4}(\widehat{\sigma}^2 - \sigma^2)^2 \leq \kappa_\alpha\right\}$$

is an approximate α-confidence region of θ^*, i.e., $\boldsymbol{\Theta}_t$ is such that

$$\mathbb{P}_{\theta^*}\left(\theta^* \in \boldsymbol{\Theta}_t\right) \approx 1 - \alpha$$

[See Bielecki et al. (2016) for more details]

Example : uncertain log-normal model

The adaptive robust control problem can be solved using the following dynamic programming principle :

$$W_T(x) = \ell\left[(\Phi(s) - v)^+\right], \quad x = (s, v, \widehat{\mu}, \widehat{\sigma}^2) \in E_X,$$

$$W_t(x) = \inf_{a \in A} \sup_{(\mu, \sigma^2) \in \tau(t, \widehat{\mu}, \widehat{\sigma}^2)} \int_{E_X} W_{t+1}(y) Q(dy \mid t, x, a; \mu, \sigma^2)$$

where $x = (s, v, \widehat{\mu}, \widehat{\sigma}^2) \in E_X = \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+$, $t = T - 1, \ldots, 0$

Example : uncertain log-normal model

The integral in the previous slide can be written as

$$\int_{\mathbb{R}} W_{t+1}\left(se^{\mu+\sigma z}, v + as(e^{\mu+\sigma z} - 1), R(t, \widehat{\mu}, \widehat{\sigma}^2, \mu + \sigma z)\right) \phi(z)dz$$

where ϕ is the density of the standard normal distribution and R is such that

$$R\left(t, \widehat{\mu}, \widehat{\sigma}^2, y\right) = \left(\frac{t}{t+1}\widehat{\mu} + \frac{1}{t+1}y, \frac{t}{t+1}\widehat{\sigma}^2 + \frac{t}{(t+1)^2}(\widehat{\mu} - y)^2\right)$$

Perspectives

- Numerically solve Bellman equation for the considered hedging problem : challenging issue due to the curse of dimensionality (optimal quantization, approximate dynamic programming could be used)
- Compare hedging performance with other approaches : control without uncertainty, standard robust, adaptive robust, Bayesian adaptive robust

Thanks for your attention.

References I

Bachoc, F. (2013).
Cross Validation and Maximum Likelihood estimations of hyper-parameters of Gaussian processes with model misspecification.
Computational Statistics & Data Analysis, 66(0) :55 – 69.

Başar, T. and Bernhard, P. (1995).
H^∞-optimal control and related minimax design problems.
Systems & Control : Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, second edition.
A dynamic game approach.

Bay, X., Grammont, L., and Maatouk, H. (2016).
Generalization of the Kimeldorf-Wahba correspondence for constrained interpolation.

Bayraktar, E., Cosso, A., and Pham, H. (2014).
Robust feedback switching control : dynamic programming and viscosity solutions.
http ://arxiv.org/pdf/1409.6233v1.pdf.

Bielecki, T., Chen, T., and Cialenco, I. (2016).
Recursive Construction of Confdence Regions.
Preprint.

Chen, H. F. and Guo, L. (1991).
Identification and stochastic adaptive control.
Systems & Control : Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA.

References II

Corsi, M., Pham, H., and Runggaldier, W. (2007).
Numerical approximation by quantization of control problems in finance under partial observations.
preprint.

Cousin, A., Maatouk, H., and Rullière, D. (2015).
Kriging of Financial Term-Structures.

Cousin, A. and Niang, I. (2014).
On the Range of Admissible Term-Structures.

Golchi, S., Bingham, D., Chipman, H., and Campbell, D. (2015).
Monotone Emulation of Computer Experiments.
SIAM/ASA Journal on Uncertainty Quantification, 3(1) :370–392.

Hansen, L. P., Sargent, T. J., Turmuhambetova, G., and Williams, N. (2006).
Robust control and model misspecification.
J. Econom. Theory, 128(1) :45–90.

Hansen, P. L. and Sargent, T. J. (2008).
Robustness.
Princeton University Press.

References III

Iyengar, G. N. (2005).
Robust dynamic programming.
Mathematics of Operations Research, 30(2) :pp. 257–280.

Kumar, P. R. and Varaiya, P. (1986).
Stochastic systems : estimation, identification and adaptive control.
Prentice-Hall, Inc.

Maatouk, H. and Bay, X. (2014a).
A New Rejection Sampling Method for Truncated Multivariate Gaussian Random Variables Restricted to Convex Sets.
To appear in Monte Carlo and Quasi-Monte Carlo Methods 2014, Springer-Verlag, Berlin 2016.

Maatouk, H. and Bay, X. (2014b).
Gaussian Process Emulators for Computer Experiments with Inequality Constraints.
in revision SIAM/ASA J. Uncertainty Quantification.

Runggaldier, W., Trivellato, B., and Vargiolu, T. (2002).
A Bayesian adaptive control approach to risk management in a binomial model.
In *Seminar on Stochastic Analysis, Random Fields and Applications, III (Ascona, 1999)*, volume 52 of *Progr. Probab.*, pages 243–258.

Sirbu, M. (2014).
A note on the strong formulation of stochastic control problems with model uncertainty.
Electronic Communications in Probability, 19.

References IV