

Credit risk and incomplete information: linear filtering and EM parameter estimation

Wolfgang Runggaldier
joint with Claudio Fontana

Dipartimento di Matematica Pura ed Applicata, Università di Padova

Evry, June 2008

Setting

An intensity-based model in an incomplete information framework

Interest rate and default intensity as affine functions of a **stochastic factor process that is not fully observable.**



Information-driven default contagion

Dynamic generalization of the **frailty-based approach**
(Schönbucher-03, Duffie et al.-06, Azizpour, Giesecke-08)

Goals and challenges

- Main goal: **valuation of illiquid credit risky products** (*corporate bonds, Credit Default Swaps,...*), for which the values can be expressed as functions of the factors (pricing measure)
- Additional goal: computation of (real-world) **default probabilities** and other quantities related to **risk management**.

Problems to be solved:

- *The factors are not observed directly*
- *The calibration of the parameters that characterize the model is not straightforward in this incomplete information setting*

Main contribution

An unified methodology for dealing with both these problems

- based on linear filtering
 - leading to a valuation procedure that is coherent with market data
 - avoiding standard calibration techniques (*that may lack stability*)
-
- Parameter estimation via *Expectation Maximization (EM)* algorithm, which is **well suited to the filtering framework**
 - The parameters are continuously updated, thus allowing the model to **track the actual market situation**

Outline

- The underlying model
- The incomplete information framework
- The filtering problem
- Parameter estimation via *EM* algorithm
- (*Risk premia* and *rating-based* information)

Intensity-based model

A framework with a number m of defaultable firms

- τ_j : random time of default of firm j , $j = 1, \dots, m$;
- $H_t^j := \mathbf{1}_{\{\tau_j \leq t\}}$; $H_t := (H_t^1, \dots, H_t^m)$;
- $\mathcal{H}_t = \sigma(H_s, s \leq t)$
- $(\mathcal{F}_t)_{0 \leq t \leq T}$ a given *background filtration*. The underlying filtered probability space is $(\Omega, \mathcal{G}, \mathcal{G}_t, P)$ with $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$, $t \in [0, T]$, $\mathcal{G}_T = \mathcal{G}$ (*full information filtration*);
- P martingale (pricing) measure, *numeraire*:

$$B(t) = B(0) \exp \left[\int_0^t r_s ds \right]$$
- Assume τ_j *conditionally independent, doubly stochastic random times* w.r.t. \mathcal{F}_t and with *intensities* λ_t^j given by the \mathcal{F}_t -*conditional hazard rate process* (\mathcal{F}_t -*intensity*) of τ_j

Intensity-based model

The *default intensity* process

In fact, according to the definition, for $\lambda_t^j \in \mathcal{F}_t$ and for independent exponential random variables ξ^j (parameter = 1 and independent of \mathcal{F}_∞) one has

$$\tau_j = \inf \left\{ t \geq 0 : \int_0^t \lambda_s^j ds \geq \xi^j \right\}.$$

Consequently (with $T > t$)

$$P\{\tau_j > T | \mathcal{G}_t\} = \mathbf{1}_{\{\tau_j > t\}} \mathbb{E} \left[e^{-\int_t^T \lambda_s^j ds} | \mathcal{F}_t \right]$$

and

$$P\{\tau_j \in (t, T] | \mathcal{G}_t\} \underset{(\tau \rightarrow t^+)}{\approx} \lambda_t^j (T - t)$$

Affine credit risk model

Interest rate and default intensity as linear functions of a stochastic factor process

- Consider a **stochastic factor process** $\Psi_t \in \mathbb{R}^n$ such that:

$$d\Psi_t = (C - A\Psi_t) dt + Bdw_t$$

where w_t is an (\mathcal{F}_t, P) -Wiener process.

- Define *interest rate* (r_t) and *default intensity* (λ_t^j) as:

$$\begin{cases} r_t = a + b\Psi_t \\ \lambda_t^j = c^j + d^j\Psi_t \quad j = 1, \dots, m \end{cases}$$

Correlation between interest rate and default intensity is a desirable property for a credit risk model.

Affine credit risk model; Bond prices

Default-free 0-coupon bond

$$\Pi_{DF}(t, T) = \mathbb{E} \left[e^{-\int_t^T r_s ds} \mid \mathcal{G}_t \right] = \exp [A(t, T) - B(t, T) \Psi_t]$$

Defaultable 0-coupon 0-recovery bond

$$\begin{aligned} \Pi(t, T) &= E \left\{ e^{-\int_t^T r_s ds} \mathbf{1}_{\tau > T} \mid \mathcal{G}_t \right\} = \mathbf{1}_{\tau > t} E \left\{ e^{-\int_t^T (r_s + \lambda_s) ds} \mid \mathcal{F}_t \right\} \\ &= \mathbf{1}_{\tau > t} \exp \left[\tilde{A}(t, T) - \tilde{B}(t, T) \Psi_t \right] \end{aligned}$$

- $A(t, T), B(t, T), \tilde{A}(t, T), \tilde{B}(t, T)$ satisfy ODEs with coefficients depending on those of the factor dynamics and in $\lambda_t^j = c^j + d^j \Psi_t$.

Affine credit risk model

Bond prices

Example : for scalar Ψ_t with

$$d\Psi_t = (C - A\Psi_t) dt + Bdw_t$$

and $r_t = \Psi_t$, $\lambda_t = \lambda \Psi_t$ one has

$$\begin{cases} \tilde{B}(t, T) &= \frac{(1+\lambda)}{A} (1 - e^{-A(T-t)}) \\ \tilde{A}(t, T) &= \frac{[\tilde{B}(t, T) - (1+\lambda)(T-t)][AC - \frac{1}{2}B^2(1+\lambda)]}{A^2} - \frac{B^2\tilde{B}^2(t, T)}{4A} \end{cases}$$

Affine credit risk model

Bond prices

- More generic credit-risky products, such as *corporate bonds* and *CDS spreads*, can be expressed by means of these two basic elements.
- Viceversa, a default-free and a defaultable term structure can be reconstructed from the more liquid *corporate bonds* prices and *CDS spreads*.
- **Example: (fair) spread of a CDS** (*Recovery of Par scheme*)

$$S = \frac{(1 - \delta) \left(1 - \Pi(t, T) - \int_t^T f(t, s) \Pi(t, s) ds \right)}{\sum_{k=1}^n (t_k - t_{k-1}) \Pi(t, t_k)}$$

Affine credit risk model

General products

- Price of a corporate bond with recovery and coupons

$$\begin{aligned} \Pi^{\delta,c}(t, T) &= \sum_{i=1}^n c(t_i - t_{i-1}) \Pi(t, t_i) + \Pi(t, T) \\ &+ \delta \left(\mathbf{1}_{\{\tau > t\}} - \Pi(t, T) - \int_t^T f(t, s) \Pi(t, s) ds \right) \end{aligned}$$

$f(t, s)$ denotes the forward rate (for deterministic r_t , $f(t, s)$ is replaced by r_s).

- (Fair) spread of a CDS (payment in arrears)

$$S = \frac{(1 - \delta) \left(1 - \Pi(t, T) - \int_t^T f(t, s) \Pi(t, s) ds \right)}{\sum_{k=1}^n (t_k - t_{k-1}) \Pi(t, t_k)}$$

Affine credit risk model

Yields and credit spreads

Yield of a 0-coupon default-free bond

$$YL(t, T) := -\frac{1}{T-t} \log \Pi_{DF}(t, T) = -\frac{A(t, T)}{T-t} + \frac{B(t, T)}{T-t} \Psi_t$$

Spread of a 0-coupon, 0-recovery defaultable bond w.r. to a default-free bond (same face value and maturity)

$$\begin{aligned} CS(t, T) &:= -\frac{1}{T-t} \log \left[\frac{\Pi(t, T)}{\Pi_{DF}(t, T)} \right] \\ &= \frac{A(t, T) - \tilde{A}(t, T)}{T-t} + \frac{\tilde{B}(t, T) - B(t, T)}{T-t} \Psi_t \quad t < \tau \wedge T \end{aligned}$$

- Yields and credit spreads are **affine functions of Ψ_t** .

Incomplete information

The investor filtration

Assume, w.l.o.g., that **all components of the factor process Ψ_t are unobservable (not precisely known)**.

- However, the investor can **observe market data**, in part. the *interest rate* (proxy), a number p of *yields* and a number q of *credit spreads*.
- The default indicator process (H_t) is indirectly contained in the *credit spreads*.

Investor filtration

$$\mathcal{Y}_t = \sigma\{r_s, YL(s, T_i), CS(s, T_j) : s \leq t, i = 1, \dots, p; j = 1, \dots, q\} \vee \mathcal{H}_t$$

and thus $\mathcal{H}_t \subset \mathcal{Y}_t \subset \mathcal{G}_t$.

Incomplete information

A filter-based pricing model

- **Objective:** evaluate an *OTC* credit risky-product, the price of which under complete information is given by $\Pi(t, T; \Psi_t)$
- **Main tool:** **filter distribution** of Ψ_t with respect to the investor filtration \mathcal{Y}_t and under the pricing measure P

Price in the investor filtration

$$\hat{\Pi}(t, T) = \mathbb{E}[\Pi(t, T; \Psi_t) | \mathcal{Y}_t]$$

- $\hat{\Pi}(t, T)$ is an *arbitrage-free* price, since $r_t \in \mathcal{Y}_t$
- $\hat{\Pi}(t, T)$ is coherent with the observations of market data, since the latter are the input to the filtering problem

Incomplete information

Noise terms affecting the observations

- All observable processes are **linear functions** of the unobserved factors.
- Therefore, if $1 + p + q > n$ the values of the factors can be determined from the observations.

This setting is not very realistic: **yields and credit spreads** are reconstructed from *corporate bonds* and *CDS spreads* and, therefore, **cannot be considered as perfectly observable**.



Introduce (Gombani, Jaschke, R.-05) ℓ further unobserved factors, on which r_t and λ_t^j do not depend, but which represent **additive noise terms** affecting the observations $YL(t, T_j)$ and $CS(t, T_j)$ and such that $n + \ell > 1 + p + q$.

Incomplete information

The observation system

- Augment Ψ_t (dim. n) to Ψ_t^* (dim. $n + \ell$) by adding the ℓ noise factors to Ψ_t :

Observation system

$$\begin{cases} r_t = a + b\Psi_t \\ YL(t, T_i) = \alpha^i(t) + \beta^i(t)\Psi_t^* & i = 1, \dots, p \\ CS(t, T_j) = \gamma^j(t) + \delta^j(t)\Psi_t^* & j = 1, \dots, q \end{cases}$$

where $\alpha^i(t), \beta^i(t), \gamma^j(t), \delta^j(t)$ depend on the model parameters.

- The observation system is a linear system
- The dynamics of the unobserved and noise factors are linear and Gaussian

The filtering problem

- The unobserved factors are Ψ_t^* and all the observations are linear with respect to them
 → **The filter distribution degenerates.**
- Defining observation vector Y_t as

$$Y_t := [r_t, YL(t, T_1), \dots, YL(t, T_p), CS(t, T_1), \dots, CS(t, T_q)]$$
 the observation system can be rewritten as

$$Y_t = \mu_t + M_t \Psi_t^*$$

with M_t of full rank (*only non-redundant observations*)

- One can find a **surrogate/auxiliary state process X_t of lower dimension** (*next Proposition*) and solve equivalently the filtering problem for X_t and Y_t .

Proposition

The filtering methodology

One can choose (non-uniquely) a time-varying $(n+\ell-p-q-1, n+\ell)$ matrix L_t such that $\begin{pmatrix} L_t \\ M_t \end{pmatrix}$ is invertible. Defining $X_t := L_t \Psi_t^*$, for the pair (X_t, Y_t) , one has

$$\begin{cases} dX_t &= (R_t + F_t X_t + G_t Y_t) dt + D_t dw_t \\ dY_t &= (V_t + C_t X_t + K_t Y_t) dt + E_t dw_t \end{cases}$$

with appropriate coefficients that all depend on the parameters of the given model. This state-observation system for X_t and Y_t is of the type of a **non-degenerate, linear conditionally Gaussian system**. Furthermore, there exist time varying matrices Γ_t, Δ_t and a vector ξ_t such that

$$\Psi_t = \Gamma_t X_t + \Delta_t Y_t + \xi_t$$

The filtering methodology

- The state-observation system for X_t and Y_t leads to the following **Gaussian filter distribution**:

$$p_{X_t | \mathcal{F}_t^Y} = \mathcal{N}(X_t; m_t, P_t)$$

with m_t and P_t recursively computed via **Kalman filter**.

- \mathcal{F}_t^Y differs from \mathcal{Y}_t only at the default times and we can solve as follows the pricing problem:

$$\begin{aligned} \hat{\Pi}(t, T) &= E \{ \Pi(t, T; \Gamma_t X_t + \Delta_t Y_t + \xi_t) \mid \mathcal{F}_t^Y \} \\ &= \int \Pi(t, T; \Gamma_t x + \Delta_t Y_t + \xi_t) p_{X_t | \mathcal{F}_t^Y}(dx) \end{aligned}$$

- Moreover, since 0-coupon bond prices $\Pi(t, T; \Psi_t)$ are exponentially affine in Ψ_t , we can obtain $\hat{\Pi}(t, T)$ directly from the **conditional moment generating function of X_t** .

Parameter estimation and *EM* algorithm

General description of the *EM* algorithm

- Let θ be the vector of the model parameters

The *EM* algorithm is based on the **iterative maximization**, w.r.t θ for a fixed θ' , of the following function Q :

$$Q(\theta, \theta') = \mathbb{E}_{\theta'} \left[\log \frac{dP^\theta}{dP^{\theta'}} \middle| \mathcal{F}_t^Y \right]$$

The *EM* algorithm iterates through the **two following steps**

- 1** (*Expectation*): compute $Q(\theta, \theta')$ for given θ' (a conditional expected value)
 - 2** (*Maximization*): maximize $Q(\theta, \theta')$ w.r.t θ
- The *maximization* step leads to a system of equations obtained by putting $\frac{\partial Q(\theta, \theta')}{\partial \theta} = 0$

Parameter estimation and *EM* algorithm

General description of the *EM* algorithm

- At every iteration one obtains non-decreasing values of the conditional *log-likelihood* function $\ell(\theta) = \log \mathbb{E}_{\bar{P}} \left[\frac{dP^\theta}{dP} \mid \mathcal{F}_t^Y \right]$. In fact (Campillo, Le Gland-89),

$$\begin{aligned} \ell(\theta) - \ell(\theta') &= \log \frac{E_{\bar{P}} \left\{ \frac{dP^\theta}{dP^{\theta'}} \cdot \frac{dP^{\theta'}}{dP} \mid \mathcal{F}_t^Y \right\}}{E_{\bar{P}} \left\{ \frac{dP^{\theta'}}{dP} \mid \mathcal{F}_t^Y \right\}} = \log E_{\theta'} \left\{ \frac{dP^\theta}{dP^{\theta'}} \mid \mathcal{F}_t^Y \right\} \\ &\geq E_{\theta'} \left\{ \log \frac{dP_\theta}{dP_{\theta'}} \mid \mathcal{F}_t^Y \right\} = Q(\theta, \theta') \end{aligned}$$

- This guarantees the **convergence** of the *EM* algorithm to stationary points of the *log-likelihood* function

Parameter estimation and *EM* algorithm

An alternating iterative *EM* algorithm

- A direct application of the *EM* algorithm to the filtering system for X_t and Y_t leads to **various difficulties**.

Our approach:

- 1 Perform a **time-discretization** of the system (actual observations are anyway in discrete time)
- 2 Apply an **alternating iterative *EM* algorithm** to several simpler parallel systems

Advantages:

- Diffusion coefficients can be easily estimated also by *EM*
- Parameter estimates as functions of the filter solution, no need for *smoothing* (Elliott, Krishnamurthy-99)

Alternating iterative *EM* algorithm: an example

- Let $n = 3, p = q = 1, \ell = 1$
- $\Psi_t = (\Psi_t^1, \Psi_t^2, \Psi_t^3)$; $\Psi_t^* = (\Psi_t, \Psi_t^4)$

$$\begin{cases} d\Psi_t^i &= (\alpha_i - \beta_i \Psi_t^i) dt + \sigma_i dW_t^i \quad (i = 1, 2, 3) \\ d\Psi_t^4 &= dW_t^4 \end{cases}$$

- $w_t = (w_t^1, w_t^2, w_t^3, w_t^4)$ Wiener with independent components.

$$\begin{cases} r_t &= \Psi_t^1 + \Psi_t^2 \\ \lambda_t &= \lambda (\Psi_t^2 + \Psi_t^3) \end{cases}$$

- Ψ_t^4 is a noise term that, multiplied by γ and ρ , affects the observations of *yield* and *credit spread* respectively
- **Parameters:** $\theta = (\alpha_1, \beta_1, \sigma_1, \dots, \alpha_3, \beta_3, \sigma_3, \lambda, \gamma, \rho)$

Alternating iterative *EM* algorithm: an example

After a **time discretization** with constant step size Δ we have, with $\varepsilon_k^i \sim \mathcal{N}(0, 1)$ i.i.d.:

$$\left\{ \begin{array}{l} \Psi_{k+1}^i = \alpha_i \Delta + (1 - \beta_i \Delta) \Psi_k^i + \sigma_i \sqrt{\Delta} \varepsilon_{k+1}^i, \quad i = 1, 2, 3 \\ r_k = \Psi_k^1 + \Psi_k^2 \\ Y_k^1 = a_k + b_k^1 \Psi_k^1 + b_k^2 \Psi_k^2 + \gamma \Psi_k^4 \quad (\text{yields}) \\ Y_k^2 = \tilde{a}_k + \tilde{b}_k^2 \Psi_k^2 + \tilde{b}_k^3 \Psi_k^3 + \rho \Psi_k^4 \quad (\text{credit spreads}) \end{array} \right.$$

- Since $n + \ell - p - q - 1 = 1$, can reduce Ψ_k to a scalar X_k
- Any of the three components of Ψ_k can be chosen as X_k and thus we obtain **three parallel linear Gaussian systems**. In each of these system consider only the observations which depend on the chosen $X_k = \Psi_k^i$.

Alternating iterative *EM* algorithm: an example

- Three systems obtained from putting $X_t = \Psi_t^i$, $i = 1, 2, 3$ respectively → **EM estimation of $\alpha_i, \beta_i, \sigma_i$**
- Consider two additional systems composed respectively by (Ψ_k^1, Y_k^1) and $(\Psi_k^2, \Psi_k^3, Y_k^2)$ → **EM estimation of γ and (λ, ρ)**
- Each of the systems is of the linear-Gaussian type and we can apply the **Kalman filter**



Instead of a global application of the (filter-based) EM algorithm to estimate all 12 parameters simultaneously, alternate the iterations between individual systems or subgroups of systems, while keeping fixed at their previously estimated values all the parameters that are not being currently estimated.

Alternating iterative *EM* algorithm: an example

Putting

$$\Psi_k^i \sim \mathcal{N}(m_k^i, V_k^i) \quad ; \quad u_k^i := \Psi_k^i - m_k^i \sim \mathcal{N}(0, V_k^i), \quad i = 1, 2, 3$$

the **three parallel systems** are

$$\begin{cases} \Psi_{k+1}^1 = \alpha_1 \Delta + (1 - \beta_1 \Delta) \Psi_k^1 + \sigma_1 \sqrt{\Delta} \varepsilon_{k+1}^1 \\ r_k = \Psi_k^1 + m_k^2 + u_k^2 \\ Y_k^1 = a_k + b_k^1 \Psi_k^1 + b_k^2 m_k^2 + b_k^3 u_k^2 + \gamma \Psi_k^4 \end{cases}$$

$$(\theta = (\alpha_1, \beta_1, \sigma_1); \quad \theta' = (\alpha_i, \beta_i, \sigma_i, \gamma, \rho, \lambda)_{i=2,3})$$

Alternating iterative *EM* algorithm: an example

$$\left\{ \begin{array}{l} \Psi_{k+1}^2 = \alpha_2 \Delta + (1 - \beta_2 \Delta) \Psi_k^2 + \sigma_2 \sqrt{\Delta} \varepsilon_{k+1}^2 \\ r_k = \Psi_k^2 + m_k^1 + u_k^1 \\ Y_k^1 = a_k + b_k^1 m_k^1 + b_k^2 \Psi_k^2 + b_k^1 u_k^1 + \gamma \Psi_k^4 \\ Y_k^2 = \tilde{a}_k + \tilde{b}_k^2 \Psi_k^2 + \tilde{b}_k^3 m_k^3 + \tilde{b}_k^3 u_k^3 + \rho \Psi_k^4 \end{array} \right.$$

$$(\theta = (\alpha_2, \beta_2, \sigma_2); \theta' = (\alpha_i, \beta_i, \sigma_i, \gamma, \rho, \lambda)_{i=1,3})$$

$$\left\{ \begin{array}{l} \Psi_{k+1}^3 = \alpha_3 \Delta + (1 - \beta_3 \Delta) \Psi_k^3 + \sigma_3 \sqrt{\Delta} \varepsilon_{k+1}^3 \\ Y_k^2 = \tilde{a}_k + \tilde{b}_k^2 m_k^2 + \tilde{b}_k^3 \Psi_k^3 + \tilde{b}_k^2 u_k^2 + \rho \Psi_k^4 \end{array} \right.$$

$$(\theta = (\alpha_3, \beta_3, \sigma_3); \theta' = (\alpha_i, \beta_i, \sigma_i, \gamma, \rho, \lambda)_{i=1,2})$$

Alternating iterative *EM* algorithm: an example

- Consider also the two additional systems (for the estimation of (γ, ρ, λ)):

$$\begin{cases} \Psi_{k+1}^1 = \alpha_1 \Delta + (1 - \beta_1 \Delta) \Psi_k^1 + \sigma_1 \sqrt{\Delta} \varepsilon_{k+1}^1 \\ Y_k^1 = a_k + b_k^2 r_k + (b_k^1 - b_k^2) \Psi_k^1 + \gamma \Psi_k^4 \end{cases}$$

$$(\theta = \gamma; \theta' = (\alpha_i, \beta_i, \sigma_i,)_{i=1,2,3})$$

$$\begin{cases} \Psi_{k+1}^2 = \alpha_2 \Delta + (1 - \beta_2 \Delta) \Psi_k^2 + \sigma_2 \sqrt{\Delta} \varepsilon_{k+1}^2 \\ \Psi_{k+1}^3 = \alpha_3 \Delta + (1 - \beta_3 \Delta) \Psi_k^3 + \sigma_3 \sqrt{\Delta} \varepsilon_{k+1}^3 \\ Y_k^2 = \tilde{a}_k + \tilde{b}_k^2 \Psi_k^2 + \tilde{b}_k^3 \Psi_k^3 + \rho \Psi_k^4 \end{cases}$$

$$(\theta = (\rho, \lambda); \theta' = (\alpha_i, \beta_i, \sigma_i,)_{i=1,2,3})$$

Alternating iterative *EM* algorithm: summary

Numerically the **most efficient way** is to alternate between the group of systems 1,2,5 and the group 3,4.

Procedure of the algorithm:

0. *Initialize the algorithm with a guess θ_0 for the entire vector θ of the parameters.*
1. *Apply in parallel on each of the systems 1, 2, 5 the EM algorithm to estimate $(\alpha_i, \beta_i, \sigma_i)_{i=1,2}$ and (ρ, λ) while keeping $(\alpha_3, \beta_3, \sigma_3, \gamma)$ fixed at their previous levels. The algorithm iterates through the two EM steps (expectation and maximization) until a stopping criterion is met, thereby producing estimates $(\hat{\alpha}_i, \hat{\beta}_i, \hat{\sigma}_i)_{i=1,2}, \hat{\rho}, \hat{\lambda}$*

Alternating iterative *EM* algorithm: summary

- 2. Apply in parallel on each of the systems 3, 4 the EM algorithm to estimate $(\alpha_3, \beta_3, \sigma_3, \gamma)$ keeping $(\hat{\alpha}_i, \hat{\beta}_i, \hat{\sigma}_i)_{i=1,2}, \hat{\rho}, \hat{\lambda}$ fixed at their previously estimated values. The algorithm iterates through the two EM steps until a stopping criterion is met, thereby producing estimates $(\hat{\alpha}_3, \hat{\beta}_3, \hat{\sigma}_3, \hat{\gamma})$.*
- 3. Reset θ at the most recently estimated levels and return to step 1. Terminate the entire algorithm as soon as a global stopping criterion is met.*

Alternating iterative *EM* algorithm

Simulation results

- Maturity $T = 5$ years, both for *default-free* and *defaultable bonds*
- $\Delta = 0.02$ (weekly observations)
- For “true” θ generate $(\psi_k^1, \psi_k^2, \psi_k^3, r_k, Y_k^1, Y_k^2)$ for $k = 0, \dots, 250$.
- *Generate randomly* θ_0
- Stop individual iterations as soon as the difference between successive values of all the parameters $< 10^{-5}$ up to a maximum of 1000.

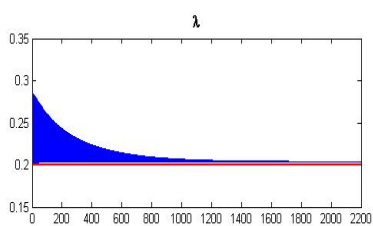
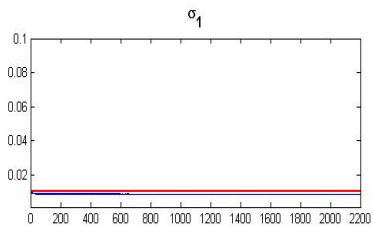
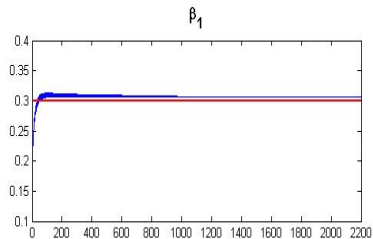
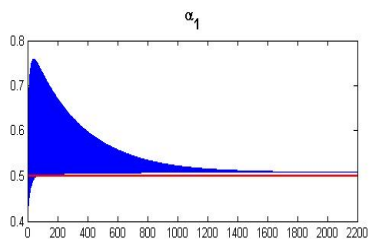
Alternating iterative *EM* algorithm

Simulation results: parameter estimates

Parameter	True value	Initial condition	Estimate
α_1	0.5	0.5870	0.5082
β_1	0.3	0.1588	0.3091
σ_1	0.01	0.1282	0.0081
α_2	0.5	0.6374	0.5082
β_2	0.3	0.4136	0.3089
σ_2	0.01	0.1778	0.0081
α_3	0.5	0.6583	0.5009
β_3	0.3	0.5019	0.3412
σ_3	0.01	0.0703	0.0218
γ	0.05	0.0453	0.0518
ρ	0.05	0.1530	0.0527
λ	0.2	0.2674	0.2070

Alternating iterative *EM* algorithm

Simulation results: convergence of the parameters $\alpha_1, \beta_1, \sigma_1, \lambda$



An extension of the model

Risk premia as further unobserved factors and *rating*-based information

- Consider the model under the **physical-historical probability measure**.
- Specify the **risk-premia**, which characterize the change of measure, as further unobserved stochastic processes to be included in the filtering system.
- In this setting we can consider also the information coming from the **rating scores**, which represent historical information.
- We can compute **filtered estimates of default probabilities**, on the basis of the information deriving from both the financial market and the *rating score*.