Operations Research Letters 38 (2010) 527–532

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Application of nonlinear filtering to credit risk

Vivek S. Borkar [a], Mrinal K. Ghosh [b,*], G. Rangarajan [b]

[a] School of Technology and Computer Science, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005, India

[b] Department of Mathematics, Indian Institute of Science, Bangalore 560 012, India

ARTICLE INFO

Article history:
Received 5 April 2010
Accepted 26 August 2010
Available online 17 September 2010

Keywords:
Merton's model
Asset
Equity
Nonlinear filter
EM algorithm

ABSTRACT

Merton's model views equity as a call option on the asset of the firm. Thus the asset is partially observed through the equity. Then using nonlinear filtering an explicit expression for likelihood ratio for underlying parameters in terms of the nonlinear filter is obtained. As the evolution of the filter itself depends on the parameters in question, this does not permit direct maximum likelihood estimation, but does pave the way for the 'Expectation–Maximization' method for estimating parameters.

1. Introduction

The assessment of credit risk is one of the most important problems in quantitative finance. A powerful approach to this is based on the option theoretic interpretation by Merton [27] (also see [4]). This approach is referred to as the *Asset Value Model* (AVM), or *structural approach to credit risk* [2,5,14,16]. While theoretically very gratifying, this still leaves wide open several computational issues. The main aim of this article is to propose a computational scheme for credit risk evaluation based on AVM. While for purposes of exposition we stick to a simple model, the underlying philosophy is broader and can be extended to more elaborate models. It has the advantage of having a rigorous footing based on methodologies that have already been utilized extensively in the signal processing community and it accounts for aspects not addressed hitherto in existing literature, as will become apparent. There is another approach to credit risk known as the reduced form (or intensity based) approach where the reason behind a default is not investigated. Instead, the dynamics of default are exogenously given through a default rate or intensity; see [1,14,16] and the references therein. We do not follow this approach in this paper.

We begin by recalling in some detail the AVM model and the current status of this problem. In this approach, the asset value process $\{A_t\}$ of the firm is assumed to follow a geometric Brownian motion (GBM) given by

$$dA_t = \nu A_t dt + \sigma A_t dW_t, \quad t \geq 0 \tag{1.1}$$

where ν is the net mean return rate on the assets, i.e., $\nu = \mu - \gamma$, where μ is the gross mean return on the assets and γ is the proportional cash payout rate; σ is the volatility, and $\{W_t\}$ is the standard Brownian motion. It is also assumed that the company has a simple *capital structure* consisting of one *debt obligation* and one type of *equity*. Let $\mathcal{E}_t$ denote the equity process of the company which is traded publicly. Suppose the process $\mathcal{D}_t$ denotes the market value of the debt obligation of the company which is assumed to have the *cash profile* of a *zero-coupon bond* maturing at a prescribed future time T and interest adjusted *face value* K. In the classical model [27] the company defaults if $A_T < K$. If the company defaults, then the payoff to the equity holders is zero. If it does not, i.e., $A_T \geq K$, then there is a net profit of $A_T - K$ after paying back the debt. Thus the total payoff to equity holders is $(A_T - K)^+ \stackrel{\text{def}}{=} \max(A_T - K, 0)$, which is identical to the payoff for a European call option on $\{A_t\}$ with strike price K, constant dividend rate γ and maturity T. Therefore for $t \in [0, T]$, $\mathcal{E}_t$ is a long European call C_t^γ from the point of view of the equity holders. Thus by the Black–Scholes–Merton option pricing formula it follows that

$$\mathcal{E}_t = C_t^\gamma = e^{-\gamma(T-t)} A_t \Phi(d_1(A_t, T-t)) - K e^{-r(T-t)} \Phi(d_2(A_t, T-t)) \tag{1.2}$$

where, for $r \stackrel{\text{def}}{=}$ the risk-free interest rate,

$$d_1(x, t) \stackrel{\text{def}}{=} \frac{\log\left(\frac{x}{K}\right) + \left(r - \gamma + \frac{1}{2}\sigma^2\right) t}{\sigma\sqrt{t}} \tag{1.3}$$

* Corresponding author.
E-mail addresses: borkar@tifr.res.in (V.S. Borkar), mkg@math.iisc.ernet.in (M.K. Ghosh), rangaraj@math.iisc.ernet.in (G. Rangarajan).

0167-6377/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.orl.2010.08.013

$$d_2(x,t) \stackrel{\text{def}}{=} \frac{\log\left(\frac{x}{K}\right) + \left(r - \gamma - \frac{1}{2}\sigma^2\right) t}{\sigma\sqrt{t}}, \tag{1.4}$$

and $\Phi(\cdot)$ as usual denotes the Gaussian distribution function. The value of the debt obligation D_T at time T is given by

$$\mathcal{D}_T = \min(K, A_T) = K - (K - A_T)^+.$$

The above payoff is equivalent to that of a portfolio consisting of a default-free loan with face value K maturing at T and a short European put option on $\{A_t\}$ with strike price K and maturity T. Thus the value of $\mathcal{D}_t$ at time t is given by

$$\mathcal{D}_t = K e^{-r(T-t)} - P_t^\gamma \tag{1.5}$$

where P_t^γ denotes the price of the put option on A_t with strike price K, constant dividend rate γ and maturity T. Using the *put-call parity*

$$A_t e^{-\gamma(T-t)} + P_t^\gamma = C_t^\gamma + K e^{-r(T-t)}$$

we obtain

$$\mathcal{D}_t = A_t e^{-\gamma(T-t)} - \mathcal{E}_t \tag{1.6}$$

where A_t and $\mathcal{E}_t$ are determined from (1.1) and (1.2) respectively. The Eq. (1.6) gives the '*theoretical*' price of the debt at time t.

Another key concept in the AVM is the default probability. In the classical model the conditional probability of default is given by

$$P(A_T < K \mid A_t) = \Phi\left(\frac{\log L_t - m(T-t)}{\sigma\sqrt{T-t}}\right), \tag{1.7}$$

where $m = \nu - \frac{1}{2}\sigma^2$ and $L_t = \frac{K}{A_t}$ is the leverage ratio of the firm at time t.

We have thus far considered the case when default occurs only at the time T of maturity of the debt. Black and Cox [3] introduced the concept of *first passage time* to compute the default probability. In this model, default occurs at a random time $\tau \in (0, T]$ when the asset value A_t falls below a level D for the first time. We assume that $D \le K$. If $D > K$, then the debt holders are fully protected [16]. More precisely, let

$$\tau_1 = \begin{cases} T & \text{if } A_T < K \\ \infty & \text{otherwise.} \end{cases}$$

Let τ_2 be the stopping time given by

$$\tau_2 = \inf\{t \in (0, T] \mid A_t < D\}.$$

Then the default time τ is given by

$$\tau = \tau_1 \wedge \tau_2.$$

Thus the forward conditional default probability at time t is given by

$$p_d(A_t) = 1 - P(\tau_1 \wedge \tau_2 > T | A_t).$$

A simple computation shows that

$$p_d(A_t) = \Phi\left(\frac{\log L_t - m(T-t)}{\sigma\sqrt{T-t}}\right) + \left(\frac{D}{A_t}\right)^{\frac{2m}{\sigma^2}} \Phi\left(\frac{\log(D^2/(KA_t)) + m(T-t)}{\sigma\sqrt{T-t}}\right). \tag{1.8}$$

This default is obviously higher than the corresponding default probability in the classical approach. Note that (1.7) is obtained as a special case of (1.8) with $D = 0$.

In the first passage model the payoff to equity holders at maturity is given by

$$\mathcal{E}_T = (A_T - K)^+ I\{M_T \ge D\}$$

where $M_t = \min_{s \le t} A_s$. The above payoff corresponds to a European down-and-out call on A_t with strike price K, barrier $D(< K)$, constant dividend rate γ and maturity T. Thus at an earlier time t, $\mathcal{E}_t$ is given by

$$\mathcal{E}_t = C_t^\gamma - e^{-\gamma(T-t)} A_t \left(\frac{D}{A_t}\right)^{\frac{2(r-\gamma)}{\sigma^2}+1} \Phi(d_3(A_t, T-t)) + K e^{-r(T-t)} \left(\frac{D}{A_t}\right)^{\frac{2(r-\gamma)}{\sigma^2}-1} \Phi(d_4(A_t, T-t)) \tag{1.9}$$

where

$$d_3(x,t) \stackrel{\text{def}}{=} \frac{\log\left(\frac{D^2}{Kx}\right) + \left(r - \gamma + \frac{1}{2}\sigma^2\right) t}{\sigma\sqrt{t}} \tag{1.10}$$

$$d_4(x,t) \stackrel{\text{def}}{=} \frac{\log\left(\frac{D^2}{Kx}\right) + \left(r - \gamma - \frac{1}{2}\sigma^2\right) t}{\sigma\sqrt{t}}. \tag{1.11}$$

In this model, the value of the debt obligation $\mathcal{D}_T$ at time T is given by

$$\mathcal{D}_T = K - (K - A_T)^+ + (A_T - K)^+ I\{M_T < D\}$$

which is equivalent to a portfolio consisting of a risk free loan with face value K, a short European put on A_t with strike price K, constant dividend rate γ and maturity T, and a long European down-in-call on A_t with strike price K, dividend rate γ, barrier D and maturity T. Therefore at an earlier time the value of the debt $\mathcal{D}_t$ is given by

$$\mathcal{D}_t = A_t e^{-\gamma(T-t)} - C_t^\gamma + e^{-\gamma(T-t)} A_t \left(\frac{D}{A_t}\right)^{\frac{2(r-\gamma)}{\sigma^2}+1} \times \Phi(d_3(A_t, T-t)) - K e^{-r(T-t)} \left(\frac{D}{A_t}\right)^{\frac{2(r-\gamma)}{\sigma^2}-1} \times \Phi(d_4(A_t, T-t)). \tag{1.12}$$

Various extensions of the first passage time models have been studied in the literature which in particular include the case when the default boundary is given by a suitable stochastic process; see [2] and the references therein. The tractability of more general models declines rapidly with growing enrichment of the models, as pointed out in [13]. The AVM is the theoretical basis for the popular commercial *estimated default frequency* (EDF) by KMV, *default probabilities* by Moody's and related ratings—see [14,22]. But these are based on historical data used in their commercial software. These procedures are proprietary and not available in the public domain. The option theoretic AVM models have also become an integral part of valuations of corporate debts using (1.2) and (1.5). One of the major difficulties in this approach is that the asset value process $\{A_t\}$ is not observable and the parameters ν, σ are unknown. Since the equity process $\{\mathcal{E}_t\}$ is traded in the market, it is therefore observable. Suppose we assume that $\{\mathcal{E}_t\}$ is also a GBM given by, say,

$$d\mathcal{E}_t = \mu_E \mathcal{E}_t dt + \sigma_E \mathcal{E}_t dW'_t \tag{1.13}$$

where $\{W'_t\}$ is a standard Brownian motion. Since $\{\mathcal{E}_t\}$ is observable, the parameters μ_E, σ_E can be estimated from the market data. Assuming $\gamma = 0$, since $\{\mathcal{E}_t\}$ is a call option on $\{A_t\}$, using Ito's formula and some additional analysis, it has been shown in [5] that

$$\frac{\sigma_E}{\sigma} = \frac{A_t}{\mathcal{E}_t} \Phi(d_1(A_t, T-t)). \tag{1.14}$$

Now A_t and σ are determined from the Eqs. (1.2) and (1.14). This is the standard textbook approach to valuing corporate debt in AVM [5] and has been studied extensively—see [2,5,14], and the references therein. Several authors have done important extensions of the AVM model, e.g., [3,15,18,20,24,25,31]. Note that by using the implicit function theorem, the Eqs. (1.2) and (1.14) can only be solved locally. These local solutions may not patch up to form a global solution for the entire planning horizon $[0, T]$. Moreover these two equations do not determine the parameter ν. Thus the default probability cannot be estimated using these equations. If we assume that entire dynamics is described under a risk neutral probability then the risk neutral default probability can "in principle" be determined using certain risk neutral instruments. Since an option on the equity $\mathcal{E}_t$ is also traded in the market, the option on the equity can be treated as a compound option on the asset A_t. This fact has been exploited in [15,18] to estimate the desired parameters. The AVM model with incomplete accounting information is addressed in [8,13]. In [13] it is assumed that a 'noisy' observation of $\{A_t\}$ denoted by Y_t is available at finitely many discrete time points $t_1, t_2, \ldots, t_n$. Under the additional assumption that the equity is not traded in the market, the conditional distribution of A_t given $\{Y_{t_1}, \ldots, Y_{t_n}\}$ is derived using Bayes rule.

In this paper we develop a new approach to estimate the parameters ν and σ based on nonlinear filtering. To motivate this, first note that though $\{\mathcal{E}_t\}$ is theoretically a call option on $\{A_t\}$, the price at which $\mathcal{E}_t$ is traded in the market may be different from its price as per the option pricing formula due to various 'noise' factors. Thus we may view $\{\int_0^t \log(\mathcal{E}_s)ds + \text{'noise'}\}$ as a process of noisy observations, say, Y_t (to be defined in the next section) of $\{A_t\}$, whence the latter is a partially observed GBM. One then has the well developed theory of nonlinear filtering that allows us to recursively estimate the conditional law of A_t given the observed $Y_s, s \le t$, for $t \ge 0$. One can explicitly write down the likelihood function for the unknown parameters in terms of the nonlinear filter whose evolution is also dependent on these parameters. This then fits the framework of maximum likelihood estimation under partial observations, for which the celebrated *Expectation–Maximization* (EM) algorithm is a well established tool. These developments are given in the next section.

The rest of our paper is structured as follows. In Section 2 we describe AVM as a partially observed GBM. Then we obtain an exact expression for the likelihood ratio which forms the basis of EM-algorithm. Section 3 deals with the estimation of parameters involved in the AVM model using extended Kalman filter. Section 4 contains some numerical results based on the procedure developed in Section 3. We conclude our paper in Section 5 with a few remarks.

2. AVM as a partially observed GBM

In this section we formalize the model we just described in the previous section. Let the asset process be described by (1.1) defined on a complete probability space $(\Omega, \mathcal{F}, P)$. Let $T > 0$ be the planning horizon as before. Let $h(t, A_t)$ denote the price of a call option on $\{A_t\}$ given either by (1.2) or by (1.9). Let E_t be the observed equity price and let $Y_t = \log E_0 + \int_0^t \log(E_s)ds$. Then Y_t is observed. Theoretically

$$Y_t = Y_0 + \int_0^t \log h(s, A_s)ds.$$

We assume that Y_t is a noisy observation of $Y_0 + \int_0^t \log h(s, A_s)ds$. Thus we assume that

$$Y_t = Y_0 + \int_0^t \log h(s, A_s)ds + cW'_t, \quad t \ge 0, \tag{2.1}$$

where $\{W'_t\}$ is a standard Brownian motion independent of $\{W_t\}$, A_0, and $c > 0$ is a constant. Define a new probability measure P_0 on $(\Omega, \mathcal{F})$ by: Under P_0, $\{A_t\}$ is given by (1.1) as before, but $c^{-1}Y_t$, $t \in [0, T]$, is a standard Brownian motion independent of $\{W_t\}$, A_0. Let $E_0[\,\cdot\,]$ denote the expectation under P_0. By Portenko's theorem [29], it follows that

$$\frac{dP}{dP_0} \stackrel{\text{def}}{=} \exp\left(\int_0^T c^{-1}\log h(t, A_t)\, dY_t - \frac{1}{2}\int_0^T c^{-2}(\log h(t, A_t))^2 dt\right).$$

Let $\mathcal{F}_t^Y \stackrel{\text{def}}{=}$ be the completion of $\cap_{t'>t}\sigma(Y_s, s \le t')$ for $t \ge 0$. Also, introduce the notation $\eta(f)$ for $\int f d\eta$ for a function–measure pair (f, η). We now state the main theorem.

Theorem 2.1. *The likelihood ratio $\Lambda_T \stackrel{\text{def}}{=} E_0[\frac{dP}{dP_0}|\mathcal{F}_T^Y]$ is given by*

$$\Lambda_T = \exp\left(c^{-1}\int_0^T \pi_t(\log h(t, \cdot))\, dY_t - \frac{c^{-2}}{2}\int_0^T \pi_t(\log h(t, \cdot))^2 dt\right) \tag{2.2}$$

where π_t is the regular conditional law of A_t given $\mathcal{F}_t^Y$ for $t \ge 0$.

Remark 2.1. $\{\pi_t\}$ is given by Fujisaki–Kallianpur–Kunita equation of nonlinear filtering:

$$\pi_t(f) = \pi_0(f) + \int_0^t \pi_s(\mathcal{L}f)\, ds + \int_0^t (\pi_s(\log h(s, \cdot)f) - \pi_s(\log h(s, \cdot))\pi_s(f))\, d\hat{Y}_s \tag{2.3}$$

$\forall f \in C_b^2(\mathbb{R}) \stackrel{\text{def}}{=}$ twice continuously differentiable bounded functions $\mathbb{R} \to \mathbb{R}$ with bounded first and second derivatives where $\mathcal{L} \stackrel{\text{def}}{=} \nu x \frac{\partial}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2}$, and $\hat{Y}_t \stackrel{\text{def}}{=} Y_t - \int_0^t \pi_s(\log h(s, \cdot))ds$, $t \ge 0$, is the so called 'innovations' process, which is a standard Brownian motion. See [6], Section V.1, for a derivation and discussion of well-posedness. The conditional law π_t is absolutely continuous with respect to the Lebesgue measure for each sample path. Let $\phi(t, x)$ denote the corresponding density. It then follows from (2.3) that $\phi(t, x)$ satisfies

$$d\phi(t, x) = (\mathcal{L}^*\phi)(t, x)dt + \phi(t, x)(\log h(t, x) - \int \phi(t, x')\log h(t, x')dx')d\hat{Y}_s \tag{2.4}$$

where $\mathcal{L}^*$ is the formal adjoint of $\mathcal{L}$.

Proof of Theorem 2.1. Recall the non-negative measure-valued process of 'unnormalized conditional laws' $\{p_t\}$ given by

$$p_t(f) \stackrel{\text{def}}{=} E_0\left[f(A_t)\exp\left(\int_0^t c^{-1}\log h(s, A_s)\, dY_s - \frac{1}{2}\int_0^t c^{-2}(\log h(s, A_s))^2 ds\right)\Big|\, \mathcal{F}_t^Y\right]$$

for $f \in C_b(\mathbb{R})$. Then its evolution is given by the Duncan–Mortensen–Zakai equation

$$p_t(f) = \pi_0(f) + \int_0^t p_s(\mathcal{L}f)\, ds + \int_0^t p_s(\log h(s, \cdot)f)\, dY_s$$

for $f \in C_b^2(\mathbb{R}^2)$ ([6], Section V.1). In particular for $f = \mathbf{1} \stackrel{\text{def}}{=}$ the constant function identically equal to 1, we have

$$p_t(\mathbf{1}) = 1 + \int_0^t p_s(\log h(s, \cdot))\, dY_s = 1 + \int_0^t \pi_s(\log h(s, \cdot))p_s(\mathbf{1})\, dY_s \tag{2.5}$$

where we have used the fact $\pi_t(f) = \frac{p_t(f)}{p_t(\mathbf{1})}$ $\forall f \in C_b(\mathbb{R})$. Treating $\{\pi_t\}$ as a 'parameter' in (2.5), direct verification shows that

$$p_t(\mathbf{1}) = \exp\left(c^{-1}\int_0^t \pi_s(\log h(s,\cdot))\,\mathrm{d}Y_s - \frac{c^{-2}}{2}\int_0^t \pi_s(\log h(s,\cdot))^2\mathrm{d}s\right)$$

is the unique solution to (2.5). Since $\Lambda_T = p_T(\mathbf{1})$, the desired result follows. □

We shall use the foregoing to estimate the unknown parameter $\theta \overset{\text{def}}{=} [\nu, \sigma, c]$. Note that both h and $\{\pi_t\}$ depend on the parameters, the latter through the dependence of $\mathcal{L}$ and h on θ in (2.3). We shall render this dependence explicit by writing $h^\theta(\cdot), \pi_t^\theta$ henceforth. From (2.2), it suffices to maximize the log-likelihood function

$$\tilde{\lambda}_T(\theta) \overset{\text{def}}{=} \left(c^{-1}\int_0^T \pi_t^\theta(\log h^\theta(t,\cdot))\,\mathrm{d}Y_t - \frac{c^{-2}}{2}\int_0^T \pi_t^\theta(\log h^\theta(t,\cdot))^2\mathrm{d}t\right).$$

To facilitate the EM algorithm, we rewrite this as

$$\lambda_T(\theta,\theta') \overset{\text{def}}{=} \left(c^{-1}\int_0^T \pi_t^\theta(\log h^{\theta'}(t,\cdot))\,\mathrm{d}Y_t - \frac{c^{-2}}{2}\int_0^T \pi_t^\theta(\log h^{\theta'}(t,\cdot))^2\mathrm{d}t\right).$$

The EM algorithm starts with an initial guess θ_0 and at step n, does the following:

1. *Expectation (E) step:* Calculate $\{\pi_t^{\theta_n}\}$. Use it to calculate $\lambda_T(\theta_n,\cdot)$.
2. *Maximization (M) step:* Find θ_{n+1} by maximizing $\lambda_T(\theta_n,\cdot)$.

It is known that this algorithm converges, albeit possibly to a local optimum. (See [10], also [9], Section 5.3. These works also state sufficient conditions for convergence to a global optimum, but these seem difficult to verify in the present context.) This can be improved upon by using multistart, simulated annealing, etc. In our case, the E-step involves calculation of the nonlinear filter. This is in principle an infinite dimensional object, but several approximation schemes exist. To begin with, there is the classical extended Kalman smoother (EKS) [21]. There are several alternative approaches of more recent vintage. These include schemes based on discretization ([23], Section 12.7), operator splitting [19], series expansions [26,28], particle filters [12], etc. One can also consider the 'pathwise filter' [11] (see also, [6], Section V.1), which is a *deterministic* parabolic partial differential equation (i.e., one not involving a stochastic integral), wherein the observation process appears as a (random) parameter. This can be approached through standard numerical techniques for parabolic p.d.e.s.

3. Estimation of parameters using extended Kalman smoother

In this section, we estimate the parameters ν, σ, c using the Extended Kalman Smoother (EKS) in the E-step of the EM algorithm. Let $X_t \overset{\text{def}}{=} \log\left(\frac{A_t}{K}\right)$. Then the Eq. (1.1) can be written as:

$$\mathrm{d}X_t = \left(\nu - \frac{\sigma^2}{2}\right)\mathrm{d}t + \sigma\,\mathrm{d}W_t, \quad t \ge 0. \tag{3.1}$$

The noisy observation Eq. (2.1) can be written as

$$\mathrm{d}Y_t = \log(h(t, A_t))\,\mathrm{d}t + c\,\mathrm{d}W'_t, \quad t \ge 0. \tag{3.2}$$

Discretizing both these equations using a step size Δt we get

$$X_k = X_{k-1} + \left(\nu - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t}\,w_k, \qquad Y'_k = g(X_k) + c\sqrt{\Delta t}\,v_k \tag{3.3}$$

where $g_k(X_k) = \log(h(k\Delta t, X_k))\Delta t$, $Y'_k = Y_k - Y_{k-1}$, w_k and v_k are independent, N(0,1)-distributed, and finally X_0 is assumed to be Gaussian with mean $\bar{x}_0$ and variance $\sigma_{x_0}^2$. If we could observe the states $\tilde{X}_N = \{X_0, X_1, \ldots, X_N\}$ in addition to the observations $\tilde{Y}'_N = \{Y'_1, Y'_2, \ldots, Y'_N\}$, under the Gaussian assumption, the complete data likelihood can be written as [30]:

$$\begin{aligned}\log L_{X,Y'}(\theta) &= -N\log 2\pi - \log\sigma_{x_0} - \frac{(X_0-\bar{x}_0)^2}{2\sigma_{x_0}^2} - N\log(\sigma\sqrt{\Delta t}) \\ &\quad - \frac{1}{2\sigma^2\Delta t}\sum_{k=1}^N\left[X_k - X_{k-1} - \left(\nu - \frac{\sigma^2}{2}\right)\Delta t\right]^2 \\ &\quad - N\log(c\sqrt{\Delta t}) - \frac{1}{2c^2\Delta t}\sum_{k=1}^N[Y'_k - g_k(x_k)]^2.\end{aligned} \tag{3.4}$$

Given that data $\tilde{X}_N$ is missing, we use the EM algorithm to iteratively find the maximum likelihood estimates of $\theta = [\nu, \sigma, c]$ based on the incomplete data $\tilde{Y}'_N$ by successively maximizing the conditional expectation of the complete data likelihood. Let

$$Q(\theta|\theta^{(j-1)}) = E[\log L_{X,Y}(\theta)|\tilde{Y}'_N, \theta^{(j-1)}], \quad j = 1, 2, \ldots. \tag{3.5}$$

The iterative process is started with an initial guess $\theta^{(0)}$ for θ. To evaluate Q (in the E step; we have suppressed the arguments in Q for notational convenience) we need to obtain the conditional expectations of each term in $\log L_{X,Y}(\theta)$. To this end we define the following quantities

$$\hat{X}_{k|j} = E[X_k|Y'_1, \ldots, Y'_j, \theta], \qquad P_{k|j} = \mathrm{var}(X_k|Y'_1, \ldots, Y'_j, \theta), \qquad P_{k,k-1|j} = \mathrm{cov}(X_k, X_{k-1}|Y'_1, \ldots, Y'_j, \theta). \tag{3.6}$$

Moreover, we linearize $g_k(x_k)$ around $\hat{X}_{k|N}$: $g_k(X_k) \approx g(\hat{X}_{k|N}) + g'_k(X_k - \hat{X}_{k|N})$, where g'_k is the derivative of $g_k(x)$ evaluated at $x = \hat{X}_{k|N}$. After some routine algebraic manipulation, we obtain the following expression for Q in terms of the quantities defined above:

$$\begin{aligned}Q &= -N\log 2\pi - \log\sigma_{x_0} - \frac{P_{0|N} + (\hat{X}_{0|N} - \bar{x}_0)^2}{2\sigma_{x_0}^2} \\ &\quad - N\log\sigma - \frac{N}{2}\log\Delta t \\ &\quad - \frac{1}{2\sigma^2\Delta t}\sum_{k=1}^N\Bigg[P_{k|N} + \hat{X}_{k|N}^2 + P_{k-1|N} + \hat{X}_{k-1|N}^2 - 2P_{k,k-1|N} \\ &\quad - 2\hat{X}_{k|N}\hat{X}_{k-1|N} - 2\left(\nu - \frac{\sigma^2}{2}\right)(\hat{X}_{k|N} - \hat{X}_{k-1|N})\Delta t \\ &\quad + \left(\nu - \frac{\sigma^2}{2}\right)^2\Delta t^2\Bigg] \\ &\quad - N\log c - \frac{N}{2}\log\Delta t - \frac{1}{2c^2\Delta t} \\ &\quad \times \sum_{k=1}^N[(Y'_k - g_k(\hat{X}_{k|N}))^2 + (g'_k)^2 P_{k|N}].\end{aligned} \tag{3.7}$$

To compute Q we therefore need to evaluate $\hat{X}_{k|N}$, $\hat{X}_{k-1|N}$, $P_{k|N}$, $P_{k-1|N}$ and $P_{k,k-1|N}$. These are given by the Extended Kalman Smoother (EKS). EKS is obtained by making one forward pass over the data followed by one backward pass. The forward pass equations are nothing but the standard EKF equations [17]. For $k = 1, 2, \ldots, N$ we have

$$\hat{X}_{k|k-1} = \hat{X}_{k-1|k-1} + \left(\nu - \frac{\sigma^2}{2}\right)\Delta t, \qquad P_{k|k-1} = P_{k-1|k-1} + \sigma^2 \Delta t$$
$$K_k = P_{k|k-1} g'_k [(g'_k)^2 P_{k|k-1} + c^2 \Delta t]^{-1},$$
$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k [Y'_k - g_k(\hat{X}_{k|k-1})]$$
$$P_{k|k} = P_{k|k-1} - K_k g'_k P_{k|k-1}$$

with the initial conditions $\hat{X}_{0|0} = \bar{x}_0$ and $P_{0|0} = \sigma^2_{x_0}$. The backward pass equations are given by (for $k = N, N-1, \ldots, 1$):

$$J_{k-1} = P_{k-1|k-1}(P_{k|k-1})^{-1}, \quad \hat{X}_{k-1|N} = \hat{X}_{k-1|k-1} + J_{k-1}(\hat{X}_{k|N} - \hat{X}_{k|k-1}), \tag{3.8}$$

$$P_{k-1|N} = P_{k-1|k-1} + J^2_{k-1}(P_{k|N} - P_{k|k-1}) \tag{3.9}$$

with initial conditions $\hat{X}_{N|N}$ and $P_{N|N}$ obtained from the forward pass equations above. Similarly, $\hat{X}_{k-1|k-1}, \hat{X}_{k|k-1}, P_{k|k-1}$ and $P_{k-1|k-1}$ are also obtained from the forward pass equations. Finally, the lag-one covariance smoother $P_{k,k-1|N}$ is given by the following equation (for $k = N, N-1, \ldots, 2$):

$$P_{k-1,k-2|N} = P_{k-1|k-1}J_{k-2} + J_{k-1}J_{k-2}(P_{k,k-1|N} - P_{k-1|k-1}), \tag{3.10}$$

with the initial condition

$$P_{N,N-1|N} = (1 - K_N g'_N)P_{N-1|N-1}. \tag{3.11}$$

In the above EKS equations, $\nu,\ \sigma$ and c should have a superscript $(j-1)$ indicating that these are the parameter values obtained in the previous iteration. This has been suppressed for notational convenience.

EKS equations enable us to compute Q as a function of θ. This completes the E step. In the M step, Q is maximized as a function of θ to obtain the jth iterate values of θ. This can be accomplished using a nonlinear optimization routine. The updated parameters are then fed into the E step and the iterative process continues till convergence is achieved.

4. Numerical results

We test the efficacy of our method using numerical simulations. We start by simulating 100 realizations of the state space equations given in Eq. (3.3) where $h(t, A_t)$ is given by Eq. (1.2). The parameter values used are: $\nu = 0.04$, $\sigma = 0.08$, $c = 0.01$, $r = 0.06$, and $\gamma = 0.03$. We integrate the equations from $t = 0$ to $t = T = 1$ using 100 time steps. Once this is done, we assume that only Y' is observed and estimate the parameters $\nu,\ \sigma$ and c from this noisy observation using the EM method (with the E step evaluated using EKS). The mean estimated values of the parameters are found to be $\hat{\nu} = 0.03$, $\hat{\sigma} = 0.06$ and $\hat{c} = 0.01$. Note that ν and σ are underestimated by the EKS method. One should be able to do better using the more sophisticated methods listed at the end of Section 2 for the E step.

Next, we repeat the above process for the state space equations given in Eq. (3.3) where $h(t, A_t)$ is now given by Eq. (1.9). The parameter values used are: $\nu = 0.04$, $\sigma = 0.08$, $c = 0.01$, $r = 0.06$, $\gamma = 0.03$, $K = 97.0$, $A_0 = 100.0$ and $D = 60$. The mean estimated values of the parameters are now found to be $\hat{\nu} = 0.028$, $\hat{\sigma} = 0.065$ and $\hat{c} = 0.01$. Again the discrepancy can be reduced by using better estimation methods.

5. Conclusion

In this paper we have developed a nonlinear filtering method for the structural approach to credit risk by treating AVM as a partially observed GBM. For the sake of simplicity we have assumed throughout the paper that the parameters involved in the model are constants which is not always a reasonable assumption. Quite often it is preferable to replace μ, γ, σ, r by continuous functions $\mu(\cdot), \gamma(\cdot), \sigma(\cdot), r(\cdot) : [0, T] \to \mathbb{R}$, with the proviso $\sigma(\cdot), r(\cdot) > 0$. In this case, the foregoing goes through with appropriate modification. The parameters $\mu(\cdot), \gamma(\cdot), \sigma(\cdot)$ are *a priori* infinite dimensional objects, but one may represent them by parsimoniously parametrized families such as linear combinations of a small number of basis functions. The problem then reduces to estimating these finite parameter vectors, which can be addressed by the EM algorithm as above. See [7] for a particular parametrization of time-variation in drift and volatility, motivated by physical considerations.

Acknowledgement

The work of the first named author is supported in part by a grant from General Motors India Ltd. The work of the second named author is supported in part by a grant from DST project: SR/S4/MS:379/06, and in part by a grant from UGC through DSA-SAP Phase IV. The work of third named author is supported in part by a DST project; also supported in part by a grant from UGC via DSA-SAP Phase IV.

References

[1] E. Bayraktar, B. Yang, A unified framework for pricing credit and equity derivatives, September, 2008 arXiv:0712.3617v2.
[2] T.R. Bielecki, M. Rutkowski, Credit Risk: Modelling, Valuation and Hedging, Springer Verlag, Berlin-Heidelberg, 2004.
[3] F. Black, J. Cox, Valuing corporate securities: some effects of bond indenture provisions, Journal of Finance 31 (1976) 351–367.
[4] F. Black, M. Scholes, The pricing of options and corporate liabilities, Journal of Political Economy 81 (1973) 637–659.
[5] C. Bluhm, L. Overbeck, C. Wagner, An Introduction to Credit Risk Modelling, Chapman and Hall / CRC, London, 2003.
[6] V.S. Borkar, Optimal Control of Diffusion Processes, Pitman Research Notes in Mathematics no. 209, Longman Scientific and Technical, Harlow, UK, 1989.
[7] V.S. Borkar, K. Sureshkumar, McKean–Vlasov limit in portfolio optimization, Stochastic Anal. Appl. (2010) (in press).
[8] M. Crouhy, D. Galai, R. Mark, Credit risk revisited, Credit Risk Supplement, March, 1998, pp. 40–44.
[9] I. Csiszar, P.C. Shields, Information theory and statistics: a tutorial, Foundations and Trends in Communications and Information Theory 1 (2004) 417–528.
[10] I. Csiszar, G. Tusnady, Information geometry and alternating minimization procedure, Statistics and Decisions (Supplement Issue No. 1) (1984) 205–237.
[11] M.H.A. Davis, Pathwise nonlinear filtering, in stochastic systems: the mathematics of filtering and identification and applications, in: M. Hazewinkel, J.C. Willems (Eds.), NATO Advanced Study Institute Series, D. Reidel, Dordrecht, 1981, pp. 505–528.
[12] P. Del Moral, Measure-valued processes and interacting particle systems. Application to nonlinear filtering problems, The Annals of Applied Probability 8 (1998) 438–495.
[13] D. Duffie, L. Lando, The term structure of credit spreads with incomplete accounting information, Econometrica 69 (2001) 633–664.
[14] D. Duffie, K.J. Singleton, Credit Risk, Princeton Uni. Press, Princeton, N.J., 2005.
[15] R. Geske, The valuation of corporate liabilities as compound options, Journal of Financial and Quantitative Analysis 5 (1977) 541–552.
[16] K. Giesecke, Credit risk modelling and valuation: an introduction, in: D. Shimko (Ed.), Credit Risk: Models and Management, vol. 2, Riskbooks, London, 2004.
[17] S. Haykin, Adaptive Filter Theory, Pearson Education, Singapore, 2002.
[18] J. Hull, I. Nelken, A. White, Merton's model, credit risk and volatility skews, Working paper, 2006.
[19] K. Ito, B.L. Rozovskii, Approximation of the Kushner equation of nonlinear filtering, SIAM Journal of Control and Optimization 38 (2000) 893–915.
[20] J.C. Jackwirth, M. Rubinstein, Recovering probabilities from option prices, Journal of Finance 51 (1996) 1611–1631.
[21] A.H. Jazwinski, Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.
[22] S. Kealhofer, Managing default risk in portfolios of derivatives, in: Derivative Credit Risk: Advances in Measurement and Management, Risk Publications, London, 1995, pp. 49–66.

[23] H.J. Kushner, P. Dupuis, Numerical Methods for Stochastic Control Problems in Continuous Time, 2nd ed., Springer Verlag, New York, 2001.
[24] H.E. Leland, B.K. Toft, Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads, Journal of Finance 51 (1996) 987–1019.
[25] F.A. Longstaff, E. Schwartz, A simple approach to valuing risky fixed and floating debt, Journal of Finance 50 (1995) 789–819.
[26] S. Lototsky, R. Mikulevicius, B.L. Rozovskii, Nonlinear filtering revisited: a spectral approach, SIAM Journal of Control and Optimization 35 (1997) 435–461.
[27] R.M. Merton, On the pricing of corporate debt: the risk structure of interest rate, Journal of Finance 29 (1974) 449–470.
[28] R. Mikulevicius, B.L. Rozovskii, Fourier–Hermite expansions for nonlinear filtering, Theory of Probability and Applications 44 (2000) 606–612.
[29] N.I. Portenko, Diffusion processes with unbounded drift coefficient, Theory of Probability and Applications 20 (1975) 27–37.
[30] R.H. Shumway, D.S. Stoffer, Time Series Analysis and its Applications, Springer, New York, 2000.
[31] C. Yi, A. Tchernitser, T. Hurd, Randomized structural models of credit spreads, Working paper, June 2009.