Forecasting Corporate Default Probabilities with Survival Models in Affine Settings

Previsione delle probabilità di default con modelli di sopravvivenza in ambito affine

Beatrice Acciaio [1]
D.E.F.S. Università di Perugia, e-mail: beatrice.acciaio@stat.unipg.it

Paolo Bordi and Elena Stanghellini
D.E.F.S. Università di Perugia, e-mail: stanghel@stat.unipg.it, bordipaolo@yahoo.it

Riassunto: In questo lavoro studiamo la probabilità di sopravvivenza di un gruppo omogeneo di agenti economici usando il metodo della forma-ridotta e assumendo un'evoluzione affine delle intensità di default. Discutiamo sia il caso continuo che quello discreto, fornendo la stima di massima verosimiglianza dei parametri di interesse in casi particolari.

1. Introduction

We observe N economic agents from a common starting point (say $t = 0$) and until their respective default or censoring times, the censoring mechanism being independent of the occurrence of default. We assume each default time τ_i ($i = 1, \ldots, N$) to admit an intensity process $\lambda_i = \{\lambda_i(t), t \geq 0\}$, which describes the conditional expected default rate (also called hazard rate in survival analysis) of agent i. This doubly-stochastic property implies that, for any time t, the probability that agent i survives to a given future time s is

$$\mathbb{P}(\tau_i > s | \mathcal{F}_t) = 1_{\{\tau_i > t\}} \mathbb{E}[e^{- \int_t^s \lambda_i(u)du} | \mathcal{F}_t], \tag{1}$$

where $\mathcal{F}_t$ is the σ-algebra that contains all the information available at time t. Intensity-based models go back to Cox and Isham (1980), see also Artzner and Delbaen (1995) and Jarrow and Turnbull (1995) for applications in this context, as well as Lando (Lando, 2004, Chapter 5) for a survey. Here, as in a large financial literature, we consider affine processes both for their computational tractability and flexibility in capturing certain properties exhibited by many financial time series. Simple examples of use of this type of processes, notoriously applied to interest-rate modeling, are the Gaussian Ornstein-Uhlenbeck model and the Feller diffusion model. In particular, we assume the default intensities λ_i to have the following dynamics:

$$d\lambda_i(t) = k(\theta - \lambda_i(t))dt + \sigma \sqrt{\lambda_i(t)}dW(t) + dJ(t), \tag{2}$$

where W is a standard Brownian motion and J is a pure-jump process (independent of W), with jump sizes independent and exponentially distributed with mean μ, and jump times of an independent Poisson process with mean jump arrival rate l (jump times and jump sizes are also independent). This assumption leads to the explicit calculation of

---

[1] Address of correspondence: Via A.Pascoli 20 - 06123 Perugia

the probabilities in (1), up to solving some ordinary differential equations (ODEs) (see Duffie *et al.* (2000) for details). The group of agents we consider is homogeneous in the sense that, due to similar economic characteristics, all the intensities λ_i can be described the same set of parameters $(k, \theta, \sigma, \mu, l)$. We here consider both the cases of continuous and discrete-times observations. In the former one we observe the triplets $(T_i, Y_i, \lambda_i(0))$, where T_i the first-exit time (due to default or censure) and Y_i is a dummy variable which equals 1 if the observation is uncensored and zero if censored. In this case the loglikelihood is as follows

$$l = \sum_{i=1}^{N} (\alpha(T_i) + \beta(T_i)\lambda_i(0)) + \sum_{i=1}^{N} Y_i \log(-\alpha'(T_i) - \beta'(T_i)\lambda_i(0)),$$

where f_i (resp. S_i) is the density (resp. survival) function associated to τ_i, and α and β are solutions of some ODEs. On the other hand, when observations occur at fixed discrete times $t = 0, 1, 2, ...$, the loglikelihood of the data set $(T_i, Y_i, (\lambda_i(t))_{t=0}^{T_i})$ takes the form

$$l = \sum_{i=1}^{N} Y_i \ln \left(\frac{P_{iT_i}}{1 - P_{iT_i}} \right) + \sum_{i=1}^{N} \sum_{j=1}^{T_i} \ln(1 - P_{ij}), \quad \text{where}$$

$$P_{ij} := \mathbb{P}(\tau_i \leq j | \tau_i > j - 1) = 1 - e^{\alpha(1) + \beta(1)\lambda_i(j-1)}, \ \forall i = 1, ..., N, \ \forall j = 1, ... T_i.$$

This leads naturally to the loglikelihood of a binary response model with exponential cumulative distribution function as a link function. The parameters of interest of the model (i.e. α and β) can be estimated, by considering each discrete time unit of each individual as a separate observation (see Cox (1970)). It then follows that a simple way to obtain an estimation of α and β is to use the Newton-Raphson algorithm where the score vector and the Hessian matrix respectively are:

$$\mathbf{s}' = \left[-\sum_{i=1}^{N} \left(\frac{Y_i}{P_{iT_i}} - T_i \right) \qquad -\sum_{i=1}^{N} \left(\frac{Y_i \lambda_i(T_i - 1)}{P_{iT_i}} - \sum_{j=1}^{T_i} \lambda_i(j - 1) \right) \right],$$

$$\mathbf{H} = \begin{bmatrix} -\sum_{i=1}^{N} Y_i \frac{1 - P_{iT_i}}{P_{iT_i}^2} & -\sum_{i=1}^{N} Y_i \lambda_i(T_i - 1) \frac{1 - P_{iT_i}}{P_{iT_i}^2} \\ -\sum_{i=1}^{N} Y_i \lambda_i(T_i - 1) \frac{1 - P_{iT_i}}{P_{iT_i}^2} & -\sum_{i=1}^{N} Y_i \lambda_i^2(T_i - 1) \frac{1 - P_{iT_i}}{P_{iT_i}^2} \end{bmatrix}.$$

References

Artzner P. and Delbaen F. (1995) Default risk insurance and incomplete markets, *Mathematical Finance*, 5, 187–195.

Cox D. (1970) *The analysis of binary data*, Methuen.

Cox D. and Isham V. (1980) *Point processes*, Chapman and Hall.

Duffie D., Pan J. and Singleton K. (2000) Transform analysis and asset pricing for affine jump-diffusions, *Econometrica*, 68, 1343–1376.

Jarrow R. and Turnbull S. (1995) Pricing derivatives on financial securities subject to credit risk, *Journal of Finance*, 50, 53–86.

Lando D. (2004) *Credit Risk Modeling: Theory and Applications*, Princeton University Press.