Estimating Term Structure Using Nonlinear Splines: A Penalized Likelihood Approach

[1]Yoshinori Kawasaki and [2]Tomohiro Ando

[1]The Institute of Statistical Mathematics, Research Organization of Information and Systems
4–6–7 Minami-Azabu, Minato-ku, Tokyo 106–8569, Japan. E-Mail: kawasaki@ism.ac.jp
[2]Graduate School of Business Administration, Keio University,
2-1-1 Hiyoshi Honcho, Kouhoku-ku, Yokohama, Kanagawa 223-8523, Japan.

Keywords: *yield curve, penalized likelihood, nonlinear functional, smoothing parameter, information criteria*

EXTENDED ABSTRACT

The spline-based models are widely used in practice to estimate the term structure of interest rates from a set of observed coupon-bond prices. The most popular method can be traced back to McCulloch (1971). Assuming that the price of a bond is equal to the present value of its future coupon payments and redemption, cash flows are regressed on a set of basis functions to estimate discount functions. Once the discount function is estimated, the zero-coupon yield and the forward rate can be obtained by transformations of the discount function. Though this method was followed by a lot of researchers, some serious drawbacks have been reported.

The most important problem is the instability of estimated yield curves. As is widely known, the discount function $\delta(t)$, the zero coupon yield $\eta(t)$ and the instantaneous forward rate $f(t)$ are closely tied with one another by explicit relationships. McCulloch's method gives approximated discount function first, so the zero coupon yield curve and the forward rate curve can be derived once $\delta(t)$ is estimated. The problem is, however, seemingly reasonable estimate of the discount function does not always lead to acceptable shapes of yield curves, especially for the forward rate curve. Some researchers are concerned with the choice of basis functions when defining a spline function, while others question how to place knots efficiently.

The choice of basis functions and/or knot locations undoubtedly affects the estimation results. However, the present article focuses on a different point. It is considered here that instability of the estimated yield curves is caused by the ill-posed nature of the regression spline, rather than by the inappropriate choice of the basis function. By ill-posed it is meant that a model may be over-parameterized compared to the amount of sample information. Without a addressing this ill-posed nature specifically, any modification of the choice of basis functions, approximating functional forms, or knots placement may provide only minor improvements.

Throughout this article, a penalty term is added to the original log-likelihood of a yield curve model, that is, a penalized likelihood approach is adopted for this treatment. In this sense, the work of Fisher, Nychka and Zervos (1995) is the most closely related and influential to this study. Those authors fitted smoothing splines (with B-splines bases) instead of regression splines. Moreover, Fisher, Nychka and Zervos (1995) fitted smoothing spline to the zero coupon yield and the forward rate as well as to the discount function. Their simulation results suggest that the best way to estimate yield curves is to place spline bases on the forward rate curve. However, splining the forward rate or the zero coupon yield is not linear operation, hence the use of GCV or the effective number of parameters in model selection lacks its theoretical foundation.

This paper proposes a penalized likelihood approach accompanied by generalized information criteria (GIC) that determine the desired degree of smoothness of yield curves in a data-dependent way. GIC, proposed by Konishi and Kitagawa (1996), is an extension of AIC, Akaike Information Criterion. Originally AIC is proposed on the assumption that the models to be compared are estimated by the method of maximum likelihood. GIC is extended to the cases where the models are not necessarily estimated by ML. Model selection among penalized (nonlinear) regressions comes within the range of GIC. Our approach is theoretically valid even if the regression functional is nonlinear with respect to the unknown coefficients of basis functions, of which typical case is 'splining the forward rate' or 'splining the zero coupon yield' case. As will be shown in Section 2, these cases are reduced to the problems of nonlinear regression spline. The derived GICs enable us to compare the models with various choices of basis functions under different regression functional forms in a unified manner. In addition, the number of basis function can be chosen based in minimum GIC method. Monte Carlo simulations reveal that choosing the appropriate number of bases by GIC reduces MSE rather than controlling a plenty of bases by a single smoothing parameter.

1 INTRODUCTION

There have been a number of studies attempting to establish an excellent technique for estimating the term structure of interest rates from a cross-section of coupon bond prices. Under the assumption that the price of a bond is equal to the present value of its future coupon payments and redemption, McCulloch (1971) regressed cash flows on a set of basis functions to estimate discount functions. Once the discount function is estimated, the zero-coupon yield and the forward rate can be obtained by transformations of the discount function.

Although the approach adopted by McCulloch (1971, 1975) was followed by several related studies, the approach has been criticized on a number of points. The central issue has been the instability of regression spline. Hence, throughout this article, a penalty term is added to the original log-likelihood of a yield curve model, that is, a penalized likelihood approach is adopted for this treatment. In this sense, the work of Fisher, Nychka and Zervos (1995) is the most closely related and influential to this study.

One important assertion made by Fisher, Nychka and Zervos (1995) based on their simulation studies is that smoothing splines could be used to spline an arbitrary transformation of the discount function. Their simulation results suggest that the best way to estimate yield curves is to place spline bases on the forward rate curve. As will be described soon below, however, splining the forward rate or the zero coupon yield is doubly nonlinear; the regression functional is nonlinear with respect to the coefficients on the spline bases, and the basis function is also nonlinear with respect to maturity, t. For such a case, the use of GCV or the effective number of parameters lacks its theoretical foundation.

It is widely known that the discount function $\delta(t)$ and the instantaneous forward rate $f(t)$ are related by

$$f(t) = -\delta'(t)/\delta(t), \tag{1}$$

where $\delta'(t)$ is the derivative of the discount function $\delta(\cdot)$ evaluated at the point t. The zero coupon yield $\eta(t)$ is tied to the discount function $\delta(t)$ by

$$\eta(t) = -\ln(\delta(t))/t. \tag{2}$$

See for example Anderson, Breedon, Deacon and Derry (1996) for the derivations of these relationships. Hence, we do not have to start by approximating the discount function $\delta(t)$. From equations (1) and (2), it is recognized that if splines are placed on $\eta(t)$ or $f(t)$, then $\delta(t)$ will be expressed as an exponential function with an approximating function for $\eta(t)$ or $f(t)$ as its argument. That is, splining the zero coupon yield or the forward rate is equivalent to exponential splining of the discount function.

By fitting a smoothing spline with cubic B-spline bases, Fisher et al. compared all three options; splining $\delta(t)$, $\eta(t)$ and $f(t)$, and determined the roughness penalty using GCV in all three cases. From a theoretical viewpoint, however, the application of GCV is questionable except the case of splining $\delta(t)$. As point out themselves (see footnote 10 and appendix B in Fisher et al. 1995), GCV cannot be applied unless the regressor is expressed as a linear combination of basis functions. In other words, a basis function can be nonlinear in t as is usual with many nonparametric regression schemes, but the regression functional should be linear with respect to the unknown parameters for the use of GCV. Clearly this does not hold in splining $\eta(t)$ or $f(t)$. Supposing that the splined zero coupon yield $\eta_s(t)$ is expressed as $\eta_s(t) = \sum w_k \phi_k(t)$, where $\{\phi_k(t); k = 1, 2, ...\}$ is a set of spline bases with coefficients w_k, then (2) implies that the splined discount function $\delta_s(t)$ is expressed as $\delta_s(t) = \exp(-t \sum w_k \phi_k(t))$. Here, δ_s is clearly not linear in $\{w_k\}$.

Sharing the motivation of Fisher et al., the aim of the present study is to propose a theoretically valid criterion that enables us to determine the desired level of smoothness of yield curves in a data-dependent way even when the regression functional is not always linear with respect to the unknown parameters. Therefore the models considered in this article are all penalized or regularized in principle. In this treatment, the generalized information criteria (GIC) introduced by Konishi and Kitagawa (1996) are tailored to various cases. Use of the GIC also makes it possible to choose the optimal number of basis functions. This is an important feature, as allowing excess knots can lead to an undesirable shape of the forward rate function. Selection of the appropriate number of basis by an objective criterion is therefore desirable.

2 PENALIZED LIKELIHOOD APPROACH

2.1 Bond equation

Consider a set of n bonds traded on one day. Let p_α be the price of bond α, c_α be its coupon payment, which is paid at time t_1^α,...,$t_{L_\alpha}^\alpha$, let R_α be the redemption payment, and let L_α be the number of remaining payments. Following the theory of bond pricing (McCulloch, 1971), we assume that the price of a bond (plus accrued interest a_α) is equal to the present value of its future coupon payments and the redemption, i.e., for $\alpha = 1, ..., n$,

$$p_\alpha + a_\alpha = c_\alpha \sum_{k=1}^{L_\alpha} \delta(t_k^\alpha) + R_\alpha \delta(t_{L_\alpha}^\alpha) + \varepsilon_\alpha, \tag{3}$$

where $\delta(\cdot)$ is the discount function, ε_α are independent and normally distributed with mean of