

正則化非線形回帰モデルによる イールドカーブの推定

川崎 能典¹・安道 知寛²

(受付 2002 年 2 月 19 日 ; 改訂 2002 年 5 月 20 日)

要 旨

McCulloch (1971) 以降、利付債のクロスセクションデータを適当な基底関数の組に回帰させて金利の期間構造を推定するという方法は、学術・実務の両面において広く定着してきた。しかし、推定されたイールドカーブ (特にフォワード・レート) の安定性や、基底関数の個数の決定などについては今なお問題が残されており、これらに関して多くの研究・提案が行われてきている。本稿では、推定曲線の不安定さは、基底関数や節点の位置の選択以上に推定問題の不適切性に起因すると考え、より一般的な解決策を目指して正則化法によるモデリングを行う。すなわち、従来の単純な最小二乗法ベースの手法に代わって、罰則付き最尤法の利用を提案する。また、割引関数とフォワード・レートを同時に安定的に推定することを可能にするために、新たにガウス基底関数に基づく統計モデルを導入する。モデル構築に当たっては、正則化 (平滑化) パラメータ、基底関数の個数、基底関数の広がりを表すパラメータの選択が本質的となるが、本稿では一般化情報量規準 GIC の枠組みに基づいて選択規準を与える。最後に日本国債のクロスセクションデータを用いて、提案した手法を実際に適用した例を示すとともに、フォワード・レートが安定的に推定されることをブートストラップ法で確認する。

キーワード：イールドカーブ，罰則付き最尤法，ガウス型動径基底関数，一般化情報量規準。

1. はじめに

金利の期間構造を利付債のクロスセクションデータから推定する問題については、これまでのところ数多くの研究が行われている。とりわけ McCulloch (1971, 1975) は、スプラインに基づく金利期間構造の推定というひとつの流れを作り出した先駆的な研究であった。McCulloch は、債券価格が将来の利払い及び償還の割引現在価値から決まるという仮定に基づき、割引関数を 3 次スプラインで近似し、期間構造を推定している。McCulloch 以後は基底関数の選択について様々な議論がなされ、Bernstein 多項式 (Schaefer (1981))、指数型スプライン (Vasicek and Fong (1982))、B-スプライン (Steely (1991)) などの利用が提案されてきた。

しかし、いずれの方法も推定においては、節点の位置をアドホックに決めた上での最小二乗法 (つまり regression spline) が用いられているため、基底関数の個数によっては推定問題がしばしば不適切問題となり、結果的に推定が不安定となる欠点を抱えている。つまり、節点の位

¹ 統計数理研究所：〒106-8569 東京都港区南麻布 4-6-7

² 九州大学大学院 数理学府：〒812-8581 福岡市東区箱崎 6-10-1

置もさることながら、基底関数の個数選択という本質的問題が残されているのである。更に実務的な問題点として、これらの方法から得られる割引関数(あるいはスポットレート)の推定精度が満足のものでも、そこから導かれるフォワード・レートは現実に解釈のしやすいものとはならないという問題点がある。

本稿においては、スポット・レートとフォワード・レートを同時に安定的に推定するために、罰則付き最尤法を用いてイールドカーブ(利回り曲線)を推定する。これは regression spline の代わりに smoothing spline を利用すると言い換えてもよい。同様の試みを行った先行研究として、Fisher et al.(1995)が挙げられる。Fisher et al.(1995)と本稿との大きな違いは、本稿ではモデル構築の際に本質的となる基底関数の個数、平滑化パラメータの選択を情報量の観点から考察し一つのモデル評価規準を導出している点にある。(Fisher et al. は基底関数の個数の選択については明確な指標を与えず、平滑化パラメータの選択に関しては一般化クロスバリデーションの利用を推奨している。)

本稿のもう一つの特徴は、基底関数としてガウス型基底関数の利用を提案していることである。この基底関数は、階層型ニューラルネットワークの一つである動径基底関数ネットワークにおける基底関数として広く用いられている。ニューラルネットワークは脳の神経細胞の数理的なモデル化に端を発する非線形モデリング手法であるが、近年では非線形構造を有する複雑な自然現象・社会現象を説明するための一つの手法として注目を浴びており、またその実用性から様々な分野で広く研究されている(Bishop(1995); Ripley(1996); Webb(1999))。なお、本稿で提案するモデリングの特徴としては、各基底関数の広がりを一様に制御するグローバルなパラメータを導入してあり、一般化情報量規準 GIC によってその適切な選択が保証されている点が挙げられる。

本稿の構成は以下の通りである。2 節では基本的な概念を述べた後、McCulloch 以来定着したイールドカーブ推定の枠組みについて言及する。また、従来から用いられている関数型やそれに伴う論点について整理し、McCulloch の 3 次スプライン基底関数の説明もおこなう。3 節では、ガウス型基底関数を導入し、罰則付き最尤法に基づきイールドカーブ推定をおこなう。4 節では、提案した方法に基づき構成されたイールドカーブモデルの評価を行うための情報量規準を導出する。すなわち、正則化法によって推定されたモデルとデータを生成したモデルとの距離を Kullback-Leibler 情報量で測ったときの推定量を一つのモデル評価規準として与える。また 5 節では、日本国債のクロスセクションデータを用いて提案した手法を実際に適用した例を示すとともに、フォワード・レートが安定的に推定されることをブートストラップ法で確認する。

2. イールドカーブの統計的推定

金利期間構造のモデリングには二つの流れがある。一方は期間構造を規定する変数を明示的に考慮した上で、何らかの均衡理論や裁定理論に基づく資産価格の決定理論に依拠するモデリングであり、他方は期間構造を規定する要因を明示的に仮定することなく、資産価格のデータからイールドカーブに対して直接的に曲線あてはめを行うことによって期間構造を推定する方法である。本稿で取り扱うのは、後者である。このように純然と統計的な方法に限っても、金利期間構造の推定は長い歴史を持つ。この点については例えば、Anderson et al.(1996)の 2.2 節を参照されたい。

2.1 割引関数とイールドカーブ

イールドカーブを推定するのに最もよく用いられるデータは、金利スワップか国債である。

スワップレートを扱うメリットは、それが実際に観測可能な金利であり、データが離散時間かつ等間隔で得られることにある。日本のスワップ市場も、1998年以降は年限20年まで活発に取引されるようになったが、それ以前となると実際に利用可能なデータはせいぜい満期が10年あたりまでである。一方、国債を用いる理由としては、デフォルトリスクを考慮する必要がないことと、長期にわたってデータが得られることが挙げられる。以後本稿では、国債のクロスセクションデータを用いてイールドカーブを推定することのみに議論を集中させる。

現在から m 年後のある日に、名目 X 円の支払いが約束された債券があるとしよう。この名目 X 円の現在価値 PV (Present Value) は、もし m 年にわたって利率が z であると仮定できれば、 $PV = X/(1+z)^m$ と表すことができる。このときの利率 z を、満期 m 年に対するゼロクーポン・イールド、あるいはしばしばスポット・レートと呼ぶ。スポットというのは、今ただちに m 年物の債券に投資したときの利回り、という意味である。ゼロクーポン・イールドと呼ぶのは、債券投資を行ってから m 年後に償還時期を迎えるまで途中の利払い (= クーポン) が無い、ということに対応している。このような債券のことを割引債、あるいはゼロクーポン債と言う。(途中の利払いをクーポンと呼ぶのは、昔は当該利払い日に債券保有者がその債券を提示し、その際に切り離し式のクーポン部分と引き替えに利払いを受け取る仕組みだったことの名残である。) 我々が関心を寄せる金利期間構造の一つは、各満期でのゼロクーポン・イールドの値の組、あるいはそれを連続的に与える曲線である。

しかし、本稿で取り扱う日本国債をはじめとして、市場に流通している債券の多くはクーポンを伴う利付債であり、例えば米国財務省証券 (T-Bill) のように割引債の価格が直接得られるケースは希である。この場合は操作的に、異なる時点で発生するキャッシュフローを各々別の割引債であるかの如く見なし、それら仮想的な割引債の複合として利付債を考えるのが一般的である。(米国財務省証券のうち、T-Notes, T-Bond は利付債として発行されるが、各々一定の条件下で「クーポンストリップス」と「プリンシパルストリップス」という割引債に交換可能で、これらの割引債も市場で取引されている。従って、これらの「ストリップス」からイールドカーブを直接推定することが可能である。この点は査読者の指摘に負う。)

従ってここでは、クーポンの額が C 、満期日にはクーポンに加えて額面 R の償還が行われる利付債を考える。一般に先進国における国債で利払いや償還が滞る可能性は皆無に等しく、ゆえにデフォルトリスクは考慮されず、将来のキャッシュフローに関する割引現在価値の合計が、そのまま現在の国債価格になっていると考えることができる。従ってもしここで、現在から将来のキャッシュフロー発生時点までのゼロクーポン・イールドが既知であれば、 m 年後に償還を迎える債券の現在価値は

$$(2.1) \quad PV_m = \frac{C}{(1+z_1)} + \frac{C}{(1+z_2)^2} + \cdots + \frac{C+R}{(1+z_m)^m}$$

と表すことができる。実際には、クーポンの支払いは年に数度行われること(例えば v 回としよう)、経過利息あるいは未収利息 (accrued interest, 以降 a と表す) の存在、経過利息の考慮にあたっては一年を何日で換算するかの慣行が国によって違うことなどを考慮に入れなければならない。ここでは日数計算の慣行に従った上で計算された償還期限 (年単位に換算) を t_L 、償還時期を含めたキャッシュフローの発生回数を L 回とし、その発生時点を t_j ($j = 1, \dots, L$)、その時点で適用可能なスポット・レートを $z(t_j)$ 、そして債券価格を p とする。これらを用いると (2.1) 式は

$$(2.2) \quad p + a = \frac{C/v}{(1+z(t_1))^{t_1}} + \frac{C/v}{(1+z(t_2))^{t_2}} + \cdots + \frac{C/v+R}{(1+z(t_L))^{t_L}}$$

と一般化できる (2.1) 式は割引現在価値の定義であるが (2.2) 式は経過利息を考慮した債券価

格が将来のキャッシュフローの割引現在価値の合計に等しい, という仮定ないしモデルに相当していることに注意が必要である. このとき (2.2) 式右辺の各項は, 名目のキャッシュフローに対してそれを現在価値に割り戻すための割引関数 $\delta(t)$ (discount function) が掛けられているとみなすことができる. すなわち, $c = C/v$ として

$$(2.3) \quad \begin{aligned} p + a &= \delta(t_1)c + \delta(t_2)c + \cdots + \delta(t_L)(c + R) \\ &= c \sum_{\alpha=1}^L \delta(t_\alpha) + \delta(t_L)R \end{aligned}$$

と書き直すことができる. 割引関数は, 1 ドルや 100 円というある 1 単位の償還を将来受けるとして, その現在価値がいくらを示している. 割引関数は利回りという意味では直観的なわかりやすさを欠くので, 通例別の量に換算される. それが以下に述べるゼロクーポン・イールドとインプライド・フォワード・レートである.

割引関数 $\delta(t)$ が与えられると, $d_j = \delta(t_j)$ によってキャッシュフロー発生時点での割引率 d_j ($j = 1, \dots, L$) が得られる. このように離散時点で定義された d_j をディスカウント・ファクターと呼ぶことがある. このとき $d_j = (1 + z_j)^{-j}$ から逆算して

$$z_j = (1/d_j)^{1/j} - 1$$

として得られるのがゼロクーポン・イールド z_j である. これらの値の組 z_j ($j = 1, \dots, L$) はスポットレートの期間構造を表しており, 時間軸を横にプロットしてつないだ曲線や, 近似曲線を連続的に描いたものは, しばしばゼロクーポン・イールドカーブと呼ばれる.

もうひとつは, インプライド・フォワード・レート (f_j と記す) である. キャッシュフロー発生時点 t_j における割引関数の値を d_j と書くと, インプライド・フォワード・レートは

$$f_j = -\frac{d_j - d_{j-1}}{d_j}$$

と定義される. つまり, ゼロクーポン・イールド z_j が j 期までの平均的な利回りに対応しているのに対し, インプライド・フォワード・レート f_j は, $j-1$ 期から j 期までの限界的な利回りを表している.

連続時間で定義すると, 割引関数 $\delta(\cdot)$ からは瞬間フォワード・レート $f(t)$ が

$$(2.4) \quad f(t) = -\frac{\delta'(t)}{\delta(t)}$$

によって定義される. ここで $\delta'(t)$ は $\delta(\cdot)$ の微分を時点 t において評価したものである. 一方, 連続時間版のゼロクーポン・イールド $\eta(t)$ は瞬間フォワード・レートの平均(積分)として

$$(2.5) \quad \eta(t) = -\frac{\ln \delta(t)}{t}$$

と表される.

2.3 節で述べるように, 割引関数 $\delta(t)$ は区分的連続関数で構成されるので, 一旦データから $\delta(t)$ が推定されると (2.4)(2.5) 式によってそれぞれ瞬間フォワード・レートとゼロクーポン・イールドが導かれる.

ここまでの議論で, 割引関数, ゼロクーポンイールドカーブ, インプライド・フォワード・レートの 3 つには明確な対応関係があり, どれかひとつが与えられれば他のカーブが導かれることは明らかである. しかし実際には, どれひとつとして直接には観測することができないので, 資産価格のデータに基づいてどれかひとつをまず推定する必要がある.

イールドカーブ推定に関する多くの学術研究は、次節で説明する McCulloch (1971) の提示した枠組みを踏襲しており、まず割引関数 $\delta(t)$ を推定し、しかる後、利回りの観点から記述する。従って手順としては、割引関数の推定値からゼロクーポン・イールドなりフォワード・レートが導かれることとなる。

2.2 割引関数の推定法

いまある日に取り引きされた国債の価格 p_α ($\alpha = 1, \dots, n$) と対応する経過利息 a_α に関するデータが与えられたとする。現時点から満期日までに償還 (R_α) 時も含めて L 回の利払い (c_α) があるとす。このとき経過利息を考慮した国債価格が将来のキャッシュフローの割引現在価値の合計に等しい

$$(2.6) \quad p_\alpha + a_\alpha = c_\alpha \sum_{\ell=1}^L \delta(t_\ell) + R_\alpha \delta(t_L), \quad \alpha = 1, \dots, n$$

と考へ、割引関数 $\delta(t)$ は、 m 個の線形独立な基底関数 $\phi_k(t)$ ($k = 1, \dots, m$) によって

$$(2.7) \quad \delta(t) = 1 + \sum_{k=1}^m w_k \phi_k(t)$$

と記述されると仮定する。基底関数の係数 w_k が推定すべき未知パラメータである。基底関数 $\phi_k(t)$ の選択についてはこれまで様々な提案がなされており、特定の関数型については 2.3 節で説明する (2.6) 式に (2.7) 式を代入すると、

$$\begin{aligned} p_\alpha + a_\alpha &= c_\alpha \sum_{\ell=1}^L \left[1 + \sum_{k=1}^m w_k \phi_k(t_\ell) \right] + R_\alpha \left[1 + \sum_{k=1}^m w_k \phi_k(t_L) \right] \\ \Leftrightarrow p_\alpha + a_\alpha - Lc_\alpha - R_\alpha &= \sum_{k=1}^m w_k \left[c_\alpha \sum_{\ell=1}^L \phi_k(t_\ell) + R_\alpha \phi_k(t_L) \right] \end{aligned}$$

を得る。ここで

$$\begin{aligned} y_\alpha &= p_\alpha + a_\alpha - Lc_\alpha - R_\alpha \\ x_{\alpha k} &= c_\alpha \sum_{\ell=1}^L \phi_k(t_\ell) + R_\alpha \phi_k(t_L) \end{aligned}$$

と定義すれば、未知パラメータ w_k ($k = 1, \dots, m$) は二乗誤差

$$l(w_1, \dots, w_m) = \sum_{\alpha=1}^n \left(y_\alpha - \sum_{k=1}^m w_k x_{\alpha k} \right)^2$$

の最小化、いわゆる最小二乗法によって推定される。割引関数 (2.7) 式のパラメータを推定したパラメータで置き換えると、割引関数

$$\hat{\delta}(t) = 1 + \sum_{k=1}^m \hat{w}_k \phi_k(t)$$

から得られる。また割引関数と同時に、瞬間フォワード・レート $f(t)$ 、ゼロクーポンイールドカーブ $\eta(t)$ も (2.4) 式、および (2.5) 式から推定される。

しかし最小二乗法は、基底関数の個数によっては逆行列の計算が不安定となる不適切問題に陥りやすく、結果として推定曲線の変動が大きいことが知られている。更に、いかなる関数型

を採用するにせよこのような基底関数による近似アプローチが避けて通れない問題は、基底の個数を如何に合理的に決定するかである。これらの問題を解決するために 3 節以降では、罰則付き最尤法を用いてイールドカーブを推定する。その際、基底関数の個数と平滑化パラメータの選択が問題となるが、4 節において情報量の観点から一つのモデル評価規準を導出することで対処する。

2.3 McCulloch の 3 次スプライン

イールドカーブ推定に用いる基底関数として、これまで様々なものが提案されている。例えば、McCulloch (1971) は 2 次スプライン基底を用いている。これは単純な多項式近似よりは合理的な結果をもたらすが、2 次スプライン基底は区分的に定義された 2 次多項式が各節点で滑らかにつながったものであるから割引関数の 2 階微分が不連続となり、そこから導かれるフォワード・レートは 1 階微分が不連続となる。これを避ける直接的な方法はスプライン関数の次元をもう 1 次元上げて、3 次スプライン基底を利用することである (McCulloch (1975))。

McCulloch の 3 次スプライン基底の構成法は、まずデータの点在する区間を各小区間に入るデータ数がほぼ等しくなるように分割し、節点の位置 k_1, \dots, k_m を決定する。ただし、本稿では節点の位置は等間隔としている。この決定した節点に基づき基底関数は以下で構成される。 $j (< m)$ 番目の基底関数 $\phi_j(t)$ は区分的多項式

$$\phi_j(t) = \begin{cases} 0 & (t < k_{j-1}) \\ \frac{(t - k_{j-1})^3}{6(k_j - k_{j-1})} & (k_{j-1} \leq t < k_j) \\ \frac{(k_j - k_{j-1})^2}{6} + \frac{(t - k_j)(t - k_{j-1})}{2} - \frac{(t - k_j)^3}{6(k_{j+1} - k_j)} & (k_j \leq t < k_{j+1}) \\ (k_{j+1} - k_{j-1}) \left\{ \frac{2k_{j+1} - k_j - k_{j-1}}{6} - \frac{t - k_{j+1}}{2} \right\} & (k_{j+1} \leq t) \end{cases}$$

で構成され (ただし、 $j = 1$ のときは $k_{j-1} = k_j = 0$ とする)、 m 番目の基底関数は t の値にかかわらず $\phi_m(t) = t$ である。

実務的に最もよく用いられるのは、ここに説明した McCulloch の 3 次スプラインモデルと思われるが、それ以外にも Bernstein 多項式を用いるもの (Schaefer (1981))、指数型スプラインを用いるもの (Vasicek and Fong (1982))、 B -スプラインを利用するもの (Steely (1991)) などが提案されてきている。

このように McCulloch の 3 次スプライン法に対して次々と新たなモデルが提案されてきたことの背景のひとつとしては、割引関数やゼロクーポン・イールドの推定という点からモデルを選んで、そこから導かれるフォワード・レートの形が必ずしも現実的に受け入れやすい形状でないということが挙げられる。Vasicek and Fong (1982) はその意味で一つの対案と言われているが、安定したフォワード・レートが得られるという証拠に乏しく、実際第三者による検証によれば結論は否定的である (Shea (1985))。フォワード・レートの形状を理にかなったものにしたと考えるならば、Nelson and Siegel (1987) のように、少数の未知母数しか持たないパラメトリックモデルで「望ましい」形のフォワード・レートのモデリングを行うことから出発する方法もある。しかしその一方、Nelson-Siegel 法のようにフォワード・レートのモデル化から入る方法を採用した場合には、今度は往々にして割引関数が解釈に苦しむ形状を呈するというような、トレードオフ問題がしばしば起こる。そこで本稿では、罰則付き最尤法を導入した上で、更に新たな基底関数として動径基底関数、とりわけガウス型基底関数の利用を提案する。これは McCulloch の 3 次スプラインに比べると、割引関数、フォワード・レート、ゼロクーポン・

イールドの全てにわたって安定度の高い推定値が得られていることがブートストラップ法により確認できる（その詳細は第5節で報告する。）

3. モデリング

この節では、本稿独自の基底関数の定式化として、動径基底関数ネットワーク(Broomhead and Lowe(1988))の枠組みで頻繁に用いられているガウス型基底関数に基づく国債価格モデルを提案する(3.1節)。割引関数は滑らかな単調減少関数となることが期待されるので、基底関数の広がり(すなわち、割引関数の滑らかさ)を制御するグローバルなパラメータを持つ滑らかな基底を選択するのは理にかなったことである。

McCulloch以降、イールドカーブの統計的推定では多くの研究が行われてきたが、そのひとつの結論として、スプラインに代表される基底関数は、データからイールドカーブを推定するにはあまりに関数型が柔軟すぎると指摘されてきた。結果としてスプラインのあてはめの際に、例えば割引関数が必ず単調減少になるように、アドホックながらも「望ましい」制約を課すことを提案したり(Shea(1985))、あるいはスプライン関数の節点の取り方に工夫を凝らす(Lizenberger and Rolfo(1984))などの試みが行われてきた。また、Nelson and Siegel(1987)のように少数のパラメータしか持たないモデルでフォワード・レートモデル化から行うという提案も、程良い柔軟性を達成するのが難しいスプライン法に対するアンチテーゼの一つとして受け止めることができる。

しかし本稿では、推定曲線の不安定さは、基底関数や節点の位置の選択問題以上に推定問題の不適切性に起因すると考え、モデルの推定にあたっては、推定の安定性、およびモデルの複雑さを考慮した罰則付き最尤法の適用を提案する(3.2節)。なお、正則化パラメータ、基底関数の個数、基底関数の広がりを表すパラメータの選択については、4節で一般化情報量規準の枠組みに基づいて一つのモデル評価規準を与える。

3.1 ガウス型動径基底関数

ある営業日に、 n 銘柄の国債が取り引きされたとする。一般に各国債の価格 p_α ($\alpha = 1, \dots, n$) は、将来のキャッシュフロー(クーポン c_α と額面 R_α)の割引現在価値に、その時点までの経過利息 a_α を考慮したもので与えられる。ここでは割引関数 $\delta(\cdot)$ をもちいて、

$$(3.1) \quad p_\alpha + a_\alpha = c_\alpha \sum_{k=1}^{L_\alpha} \delta(t_k^\alpha) + R_\alpha \delta(t_{L_\alpha}^\alpha) + \varepsilon_\alpha, \quad \alpha = 1, \dots, n,$$

と仮定する。ただし、ノイズ ε_α は互いに独立で平均0、分散 σ^2 の正規分布に従うとする。また t_k^α は当該日を起点とした k 期目の利払い日までの期間を年単位に換算したものとし、 L_α で配当回数を表す。ここで $k = L_\alpha$ すなわち満期 $t_{L_\alpha}^\alpha$ においては利払いと償還が行われることに注意する。

割引関数 $\delta(\cdot)$ に対しては、基底関数の線形和による表現を仮定する。

$$(3.2) \quad \delta(t) = 1 + \sum_{k=1}^m w_k \phi_k(t) + w_0 = 1 + \mathbf{w}' \boldsymbol{\phi}(t),$$

ここで、 $\boldsymbol{\phi}(t) = (1, \phi_1(t), \dots, \phi_m(t))'$ は基底関数から構成される既知の基底関数ベクトル、 $\mathbf{w} = (w_0, w_1, \dots, w_m)'$ は未知のパラメータベクトルである。2.3節で述べたように、様々な基底関数のとり方があるが、本稿においてはガウス基底関数を用いることを提案する。

$$(3.3) \quad \phi_k(t) = \exp\left(-\frac{(t - \mu_k)^2}{2s^2}\right), \quad k = 1, \dots, m.$$

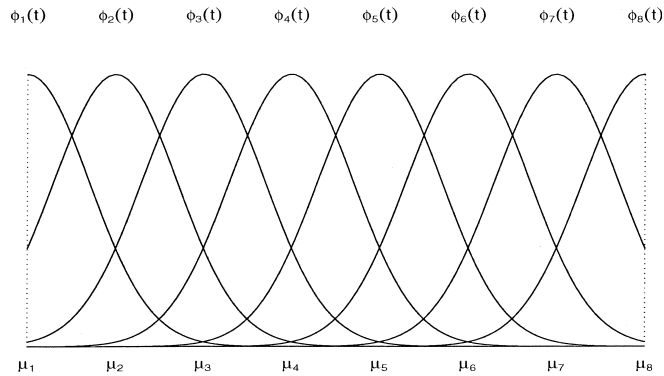


図 1. ガウス基底関数の例.

ただし, μ_k は基底関数の中心とし, s^2 は基底関数の広がりを表す量で, 割引関数の滑らかさを制御するグローバルなパラメータである. 本稿においては, データの点所在する区間を等間隔に分割し, 各小区間の節目を $\phi_k(t)$ の中心 μ_k とする. 図 1 は 8 個のガウス基底関数の例である. (3.1) 式と (3.2) 式より

$$y_\alpha = \mathbf{c}'_\alpha \Phi_\alpha \mathbf{w} + \varepsilon_\alpha, \quad (y_\alpha = p_\alpha + a_\alpha - L_\alpha c_\alpha - R_\alpha),$$

が導かれ, 国債価格モデルは

$$(3.4) \quad f(y_\alpha | t_1^\alpha, \dots, t_{L_\alpha}^\alpha; \mathbf{w}, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(y_\alpha - \mathbf{c}'_\alpha \Phi_\alpha \mathbf{w})^2}{2\sigma^2} \right\},$$

と定式化される. ただし, $\Phi_\alpha = (\phi(t_1^\alpha), \dots, \phi(t_{L_\alpha-1}^\alpha), \phi(t_{L_\alpha}^\alpha))'$, $\mathbf{c}_\alpha = (c_\alpha, \dots, c_\alpha, c_\alpha + R_\alpha)'$ とする.

次節では, 推定の安定性, およびモデルの複雑さを考慮した罰則付き最尤法に基づきモデルのパラメータ \mathbf{w}, σ^2 を推定する.

3.2 推定

ガウス基底関数に基づく国債価格モデル (3.4) 式のパラメータ \mathbf{w} を最尤法によって推定するとき, パラメータ \mathbf{w} の最尤推定値 $\hat{\mathbf{w}}$ は $\hat{\mathbf{w}} = (B'B)^{-1} B'y$ で与えられる. ただし $B = (\Phi_1' \mathbf{c}_1, \dots, \Phi_n' \mathbf{c}_n)'$, $\mathbf{y} = (y_1, \dots, y_n)'$ とする. しかし, 最尤推定値は基底関数の個数によっては逆行列の計算が不安定となり, また推定曲線の変動が大きいことが知られている. そこで, 国債価格モデル $f(y_\alpha | t_1^\alpha, \dots, t_{L_\alpha}^\alpha; \mathbf{w}, \sigma^2)$ に含まれるパラメータ \mathbf{w}, σ^2 は罰則付き対数尤度関数の最大化に基づいて推定する.

$$(3.5) \quad l_\lambda(\mathbf{w}, \sigma^2) = \sum_{\alpha=1}^n \log f(y_\alpha | t_1^\alpha, \dots, t_{L_\alpha}^\alpha; \mathbf{w}, \sigma^2) - \frac{n\lambda}{2} \mathbf{w}' K \mathbf{w}.$$

ここで, λ は正則化パラメータで, 推定の安定性を向上させる. また罰則項の $(m+1) \times (m+1)$ 差分行列 K は $\sum_{j=2}^m (\Delta^2 w_j)^2 = \mathbf{w}' K \mathbf{w}$ を表すものとする. ただし, $\Delta w_j = w_j - w_{j-1}$ である. その他の罰則項としては, 割引関数の曲率を考慮した 2 階微分の積分の離散近似, パラメータの 2 乗和等が考えられる. 罰則付き対数尤度関数 (3.5) の最大化に基づくパラメータの推定量

$\hat{w}, \hat{\sigma}^2$ は

$$\hat{w} = (B'B + n\beta K)^{-1} B'y, \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{\alpha=1}^n \{y_{\alpha} - c'_{\alpha} \Phi_{\alpha} \hat{w}\}^2.$$

で与えられる．ただし $\beta = \sigma^2 \lambda$ とする．

割引関数 (3.2) 式のパラメータ w, σ^2 を，推定したパラメータ $\hat{w}, \hat{\sigma}^2$ で置き換えると割引関数が推定される．また割引関数と同時に (2.4) 式 (2.5) 式の関係からフォワード・レート $f(t)$ ，ゼロクーポン・イールドカーブ $\eta(t)$ も推定される．

罰則付き最尤法に基づき構成した統計モデル $f(y_{\alpha}|t_1^{\alpha}, \dots, t_{L_{\alpha}}^{\alpha}; \hat{w}, \hat{\sigma}^2)$ は基底関数の個数 m ，正則化パラメータ λ ，ガウス基底関数に含まれる s^2 に依存し，従ってこれらの値を適切に選択する必要がある．次節では，統計モデル $f(y_{\alpha}|t_1^{\alpha}, \dots, t_{L_{\alpha}}^{\alpha}; \hat{w}, \hat{\sigma}^2)$ を評価するための情報量規準を導出し， m, λ, s^2 を選択する手法について述べる．

なお，本稿と同様の試みを行った先行研究として，Fisher et al. (1995) が挙げられるので，本稿との相違点についてここでコメントしておく．Fisher らは罰則付き最小二乗法 (smoothing spline) を用いているが，本稿の罰則付き最尤法も誤差分布の仮定が正規分布という意味では同一の接近法である．一方，基底関数については，Fisher 他は 3 次 B-スプラインを採用しており，本稿ではガウス型基底関数を採用している．(また，比較対照にあたっては McCulloch の 3 次スプラインも用いている．) 最も大きな違いは，本稿ではモデル構築の際に本質的となる基底関数の個数，平滑化パラメータの選択を情報量の観点から考察し一つのモデル評価規準を次節で導出している点にある．Fisher 他は基底関数の個数の選択については明確な指標を与えず，平滑化パラメータの選択のみに関して一般化クロスバリデーションの利用を推奨している．

4. 一般化情報量規準

Akaike (1974) は，想定したモデルとデータを生成した真のモデルとの距離を Kullback-Leibler 情報量 (Kullback and Leibler (1951)) で測ったとき，想定したモデルの近くに真のモデルがあれば，最尤法に基づき構成したモデルの対数尤度のバイアスは，モデルのパラメータ数で近似できるとした．しかし，想定したモデルの近くに真のモデルが常にあるとは限らないこと，AIC は最尤法に基づき構成したモデルを評価する規準であり，本稿では罰則付き最尤法に基づき統計モデル $f(y_{\alpha}|t_1^{\alpha}, \dots, t_{L_{\alpha}}^{\alpha}; \hat{w}, \hat{\sigma}^2)$ を構成していることから，単純に AIC を適用することには問題がある．

Konishi and Kitagawa (1996) は，一般に最尤法以外の方法によって推定されたモデルとデータを生成した真のモデルとの距離を Kullback-Leibler 情報量で測るという立場から，その距離の推定量を一つのモデル評価規準として与えた．この結果を用いると，その漸近的な推定量として次の情報量規準が導出される．

$$(4.1) \quad \begin{aligned} \text{GIC}(m, \lambda, s^2) &= -2 \sum_{\alpha=1}^n f(y_{\alpha}|t_1^{\alpha}, \dots, t_{L_{\alpha}}^{\alpha}; \hat{w}, \hat{\sigma}^2) + 2\text{tr}\{I_G J_G^{-1}\} \\ &= n \log(2\pi\hat{\sigma}^2) + n + 2\text{tr}\{I_G J_G^{-1}\}. \end{aligned}$$

ただし I_G と J_G は $(m+2) \times (m+2)$ でそれぞれ

$$\begin{aligned} I_G &= \frac{1}{n\hat{\sigma}^2} \begin{pmatrix} B'\Lambda/\hat{\sigma}^2 - \lambda K \hat{w} \mathbf{1}'_n \\ \mathbf{p}' \end{pmatrix} (\Lambda B, \hat{\sigma}^2 \mathbf{p}), \\ J_G &= \frac{1}{n\hat{\sigma}^2} \begin{pmatrix} B'B + n\hat{\sigma}^2 \lambda K & B'\Lambda \mathbf{1}_n / \hat{\sigma}^2 \\ \mathbf{1}'_n \Lambda B / \hat{\sigma}^2 & n/2\hat{\sigma}^2 \end{pmatrix}. \end{aligned}$$

で与えられる．ただし $\Lambda = \text{diag}[y_1 - c'_1 \Phi_1 \hat{w}, \dots, y_n - c'_n \Phi_n \hat{w}]$, $\mathbf{1}_n = (1, 1, \dots, 1)'$ とし, n 次元ベクトル p の第 i 番目の成分は $(y_i - c'_i \Phi_i \hat{w})^2 / 2\hat{\sigma}^4 - 1/2\hat{\sigma}^2$ である．この GIC を最小にする m, λ, s に対応するモデルを最適なモデルとして選択する．

また, 提案するモデル評価規準 GIC は, McCulloch の 3 次スプライン, Bernstein 多項式 (Schaefer (1981)), B -スプライン (Steely (1991)) 等を (3.2) 式の基底関数として罰則付き最尤法でモデルを推定した場合にも, 同様のモデル評価が可能である．

5. 解析例と安定度分析

5.1 事例解析

本節では, 国債取引データの分析を通して提案する手法の有効性を検証する．まず, 特定の営業日に観察された一組のデータを解析した事例として, ある程度銘柄数を確保できた 1997 年 7 月 30 日 (65 銘柄) を選んで分析した結果を報告する．

97 年 7 月 30 日に取り引きされた国債銘柄を用いて推定されたイールドカーブが図 2 に示されている．図 2 (a) (b) (c) はそれぞれ推定された割引関数, 瞬間フォワード・レート, ゼロクーポン・イールドとなっている．基底関数の個数 m と平滑化パラメータ λ の組み合わせにより (4.1) 式の GIC の値は変化するが, ここでは広範囲で格子探索を行っているので, 数表で報告する代わりに鳥瞰図 (図 2 (d)) で挙動を示している．ただし平滑化パラメータ λ については対数変換して探索しており, ガウス基底関数に含まれる s^2 は m, λ の各組み合わせにおいて最適化している．格子探索により, 最適な基底関数の個数 m 平滑化パラメータ λ はそれぞれ

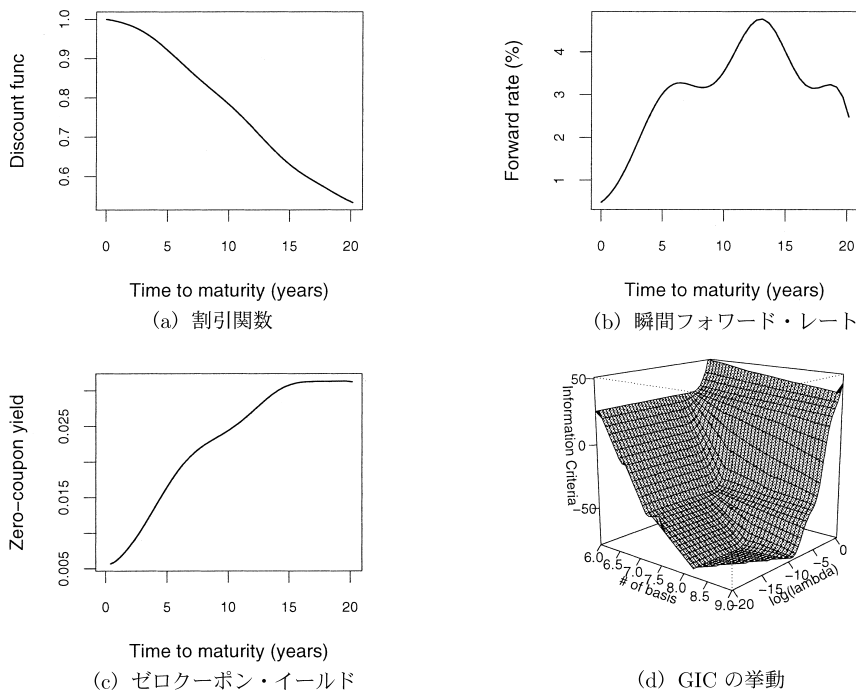


図 2. 97 年 7 月 30 日のデータから推定された (a) 割引関数 (b) 瞬間フォワード・レート, (c) ゼロクーポン・イールドと (d) 情報量規準の挙動．

$\hat{m} = 8, \hat{\lambda} = 10^{-9.3}$ という結果が得られ、またそのとき $\hat{s}^2 = 19.9$ であった。

推定されたフォワード・レートの形状については、以下のような解釈が可能である。まず、5~6年近辺のこぶは、いわゆる指標銘柄効果によって9年近辺のスポットレートが低いことが原因であると考えられる。一方、既にこのころから指標銘柄効果が薄れてきている可能性を考慮すると、先物受渡銘柄(最割安銘柄)の効果も曲線の形状に影響を与えていると考えられる。(この点は査読者の指摘に負う。)また、13~14年付近のこぶは、長期債と超長期債の流動性の違いに起因すると考えられる。長期債(発行年限10年)は取引量が多く流動性も高いが、超長期債(発行年限20年)は取引量・流動性とも長期債に比べると落ちるため、流動性プレミアムが割引かれることによって割安となる。このことから、13~14年付近にこぶが現れると解釈できる。

このように、 m, λ, s^2 の選択において、客観的な指標を用いることができるのが本稿で提案する方法の最大のメリットであるが、一方で推定されたモデルの安定性を検証することも興味ある課題である。そこで5.2節では、ブートストラップ法を適用して、本稿で提案する方法から推定されるイールドカーブ(特に瞬間フォワード・レート)の安定性を吟味し、これによって従来型手法とのパフォーマンスの差異を明らかにする。

5.2 ブートストラップ法による安定性分析

ブートストラップ法を用いて、提案する手法と2.2節で紹介した McCulloch の3次スプラインによる割引関数の推定法の安定性を比較検証する。つまり、提案するガウス基底関数モデルは罰則付き最尤法と GIC に基づいてモデリングをおこない、McCulloch の3次スプラインモデルは通常最小二乗法で推定する。また McCulloch (1971)には基底関数の個数の選択に関しては何も述べられていないことから、ここではクロスバリデーション(CV)を用いて基底関数の個数を選択した。

ここでは2001年8月22日の国債データ(78銘柄)を用いる。図3~図5は、2001年8月22日に取引された国債データから100組のブートストラップ標本を発生させ、各ブートストラップ標本に基づいてイールドカーブを推定したものである。ただし、ブートストラップ標本は、データの組を復元抽出することで構成した。(その理由は、残差の不均一性にある。図3~図5から観察されることなのだが、満期までの期間が長い区間では、推定されたイールドカーブはより大きな不確実性を抱えている。実際多くの事例で、償還期限が10年未満と10年を越えたところでは、モデルの残差分散は異なっているように見える。このような場合に残差に基づくブートストラップを行うことは、本来なら安定的に推定されているはずである償還期限の短いところでのイールドカーブの変動性を過大評価したり、逆の副作用を償還期限の長いところでもたらすと考えられる。)

図3によれば、割引関数の観点からは一見両者に大きな違いはないように思われる。しかしながら、図4(a)から明らかなように、フォワード・レートについては罰則付き最尤法に GIC を併用することにより極めて安定的に推定できることがわかる。対照的に図4(b)の McCulloch の3次スプラインモデルは、正則化なしの最小二乗法で推定されているためイールドカーブの変動が大きく、ブートストラップ標本によってはパラメータ推定が極度に不安定になる場合もあることが観察される。

イールドカーブは日々再推定し描き直すものであるが、日によっては時間軸(満期までの期間)の中にぽっかりとデータのない区間が現れることもある。満期期間が長いところでは、いずれのケースでも若干カーブが膨らんでおり、この期間での推定値が不安定であることを示している。このようにデータが疎な区間では、仮にイールドカーブを引けたとしても殆ど仮定した関数型に依存することになり、このようにブートストラップ法を適用すると、期間によって

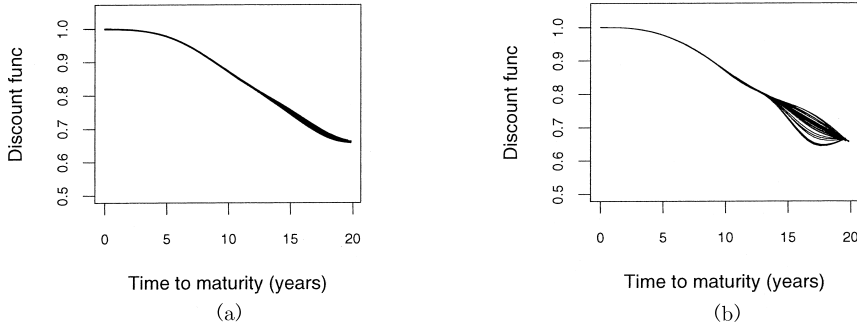


図 3. ブートストラップで推定された割引関数(2001年8月22日)(a)提案する手法, (b)McCullochの3次スプラインモデル.

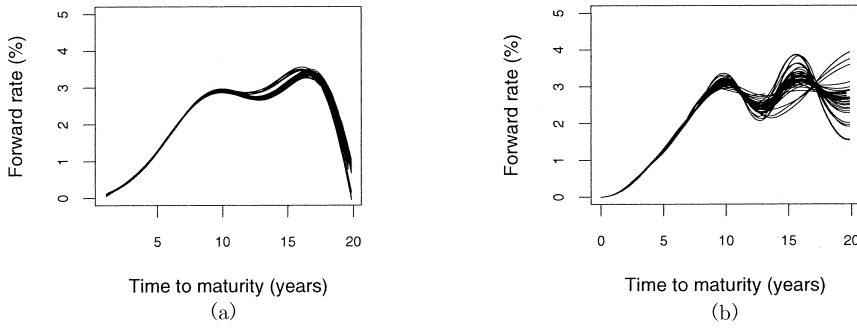


図 4. ブートストラップで推定されたフォワード・レート(2001年8月22日)(a)提案する手法(b)McCullochの3次スプラインモデル.

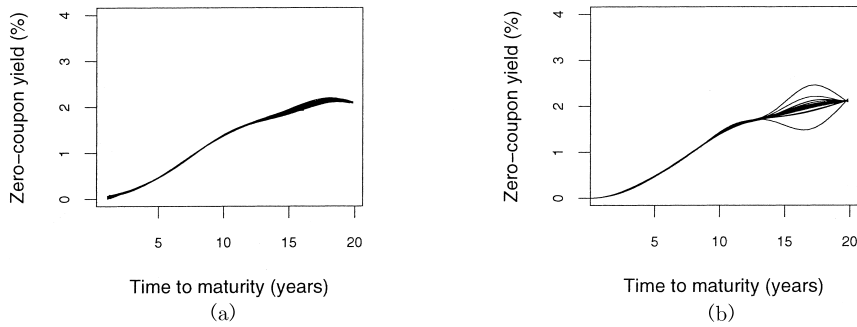


図 5. ブートストラップで推定されたゼロクーポン・イールド(2001年8月22日)(a)提案する手法(b)McCullochの3次スプラインモデル.

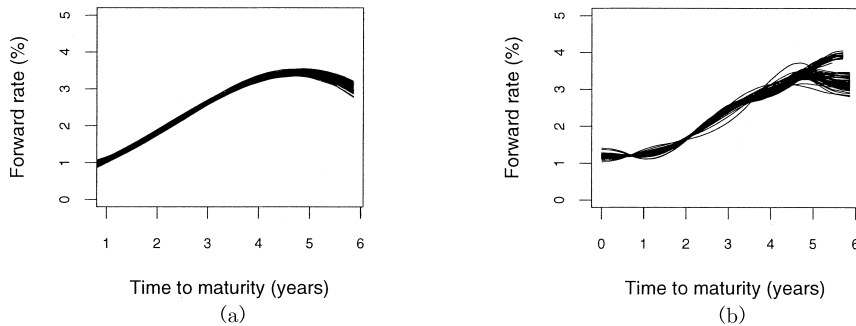


図 6. ブートストラップで推定されたフォワード・レート(95年6月14日)(a)ガウス基底関数に基づく結果 (b)McCulloch の3次スプラインモデルに基づく結果。

信頼性に差があることが確認できる。それでも、フォワード・レートのばらつきを比較すれば、本稿で提案するモデルと評価法が安定性の意味ですぐれていることは一目瞭然だろう。

このように、データが疎な区間ではモデルの構造を「堅め」にしておかなければ、オーバーフィッティングが起きる危険性は少なくない。特に基底関数の個数を合理的に決める方法を持たないときにはなおさらである。割引関数の空間では一見して精度よいあてはめが行われているようであっても、過剰適合に由来するわずかな凹凸が、フォワード・レートに変換すると大きな山や谷になって現れる危険性がある。

ここまでは、正則化項を導入して適切なモデリングを行うこと、一般化情報量規準 GIC を利用して適切なモデル選択を行うことの重要性を、実例を通じて示してきた。最後に、ガウス型基底関数を採用する利点を強調するために、1995年6月14日(22銘柄)の国債価格データを分析した結果を報告しておこう。先ほどの図3~図5の場合とは異なり、ここでは McCulloch の3次スプラインのケースにも正則化を行い、一般化情報量規準 GIC によって正則化パラメータ、基底の個数等を決めている。従って、2つの手法の比較において異なる部分は、基底関数の選択だけである。

先ほどと同様に、割引関数の観点からは一見両者に大きな違いはなかった。しかしながら、図6(a)から明らかなように、フォワード・レートについては、ガウス型動径基底関数に基づく方が推定曲線が安定している。特にこのケースは、2001年に比べると満期の長いデータが殆どないので、図6(b)の3次スプラインの結果は5年未満のところですらフォワード・レートが非常に安定性を欠いていることがわかる。これは(3.3)式における s^2 が、基底全体の堅さを制御することに成功している例である。ちなみに、97年7月30日の解析例では $\hat{s}^2 = 19.9$ であったが、95年6月14日のケースでは $\hat{s}^2 = 0.98$ であり、値が大きく違うことがわかる。このようにデータが少ない場合でも、ガウス型動径基底関数に罰則付き最尤法と情報量規準を併用することで、半自動的に最適な s^2 の選択が可能になり、適切なモデリングを行うことができる。また、紙幅の関係上多くを紹介できないが、提案する手法を様々な年度のデータに適用した結果、割引関数、フォワード・レート、ゼロクーポン・イールドの推定が極めて安定し、正則化と GIC が有効に機能していることが一般的に確認された。

6. まとめ

本稿では、利付債のクロスセクションデータからイールドカーブを推定する方法について、幾つかの新しい提案を行った。McCulloch(1971)に倣って割引関数の推定を行う際、従来の最

小二乗法ではなく、罰則付き最尤法によってモデルを推定することを提案した。更に、基底関数として新たにガウス型動径基底関数を利用するモデリングを提案した。このモデルの特徴は、各基底関数の広がりを一様に制御するグローバルなパラメータを持つ点にある。正則化パラメータ、基底関数の個数、基底関数の広がりを表すパラメータについては、一般化情報量規準 GIC の枠組みに基づいて選択規準を与えた。これによって、従来曖昧だった基底関数の個数選択問題について明確な回答を与えることができた。提案した方法により得られたイールドカーブの安定性をブートストラップ法により検証すると、割引関数、ゼロクーポン・イールドと同時にフォワード・レートも安定的であることがわかる。特に銘柄数の少ない時には、ガウス型基底関数のグローバルパラメータが重要な役割を果たすことが示された。

謝 辞

二人の査読者の丁寧なコメントのおかげで内容は大幅に改善されました。ここに記して感謝いたします。また、本研究の初期段階において、分析結果の解釈等について何度も貴重なコメントをお寄せくださった、桑原浩人氏(三菱証券)に感謝致します。

参 考 文 献

- Akaike, H. (1974). A new look at the statistical model identification, *IEEE Trans. Automat. Control*, **AC-19**, 716–723.
- Anderson, N., Breedon, F., Deacon, M., Derry, A. and Murphy, G. (1996). *Estimating and Interpreting the Yield Curve*, Wiley, Chichester.
- Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford.
- Broomhead, D. S. and Lowe, D. (1988). Multivariable functional interpolation and adaptive networks, *Complex Systems*, **2**, 321–335.
- Fisher, M. E., Nychka, D. and Zervos, D. (1995). Fitting the term structure of interest rates with smoothing splines, Federal Reserve Bank Finance and Economics Discussion Paper 95-1, January.
- Konishi, S. and Kitagawa, G. (1996). Generalised information criteria in model selection, *Biometrika*, **83**, 875–890.
- Kullback, S. and Leibler, R. A. (1951). On information and sufficiency, *Ann. Math. Statist.*, **22**, 79–86.
- Litzenberger, R. H. and Rolfo, R. (1984). An international study of tax effects on government bonds, *Journal of Finance*, **39**(1), 1–22.
- McCulloch, J. H. (1971). Measuring the term structure of interest rates, *Journal of Business*, **44**(1), 19–31.
- McCulloch, J. H. (1975). The tax-adjusted yield curve, *Journal of Finance*, **30**(3), 811–830.
- Nelson, C. R. and Siegel, A. F. (1987). Parsimonious modeling of yield curves, *Journal of Business*, **60**(4), 473–489.
- Ripley, B. D. (1996). *Pattern Recognition and Neural Networks*, Cambridge University Press, Cambridge.
- Schaefer, S. M. (1981). Measuring a tax-specific term structure of interest rates in the market for British government securities, *The Economic Journal*, **91**, 415–438.
- Shea, G. S. (1985). Interest rate term structure estimation with exponential splines, *Journal of Finance and Quantitative Analysis*, **19**, 253–269.

- Steely, J. M. (1991). Estimating the gilt-edged term structure: Basis splines and confidence intervals, *Journal of Business, Finance and Accounting*, **18**(4) (June), 512–529.
- Vasicek, O. A. and Fong, H. G. (1982). Term structure modeling using exponential splines, *Journal of Finance*, **37**(2), 339–356.
- Webb, A. (1999). *Statistical Pattern Recognition*, Arnold, London.

Nonlinear Regression Models with Regularization and Their Application to Yield Curve Estimation

Yoshinori Kawasaki

(The Institute of Statistical Mathematics)

Tomohiro Ando

(Graduate School of Mathematics, Kyushu University)

There have been a considerable number of researches on estimation of yield curves. Since McCulloch (1971), utilizing cross-sectional data of coupon bonds, a method for regressing cash flows on a set of basis functions to estimate discount factors has been widely used and discussed by both academic researchers and practitioners. Attention has focused on the stability of the implied forward rate curve, and also on the optimal choice of the number of basis functions. However, there seems to be no definitive method for overcoming these shortcomings. Based on the observation that the least squares approach tends to result in a so-called ‘improper’ problem, this article presents a regularization method or penalized maximum likelihood approach to stabilize the shapes of yield curves. It also proposes a Gaussian radial basis function as an alternative to commonly adopted basis functions such as McCulloch’s natural cubic spline. In model determination, what is essential is the choice of regularization parameter, the number of basis and the dispersion parameter in the Gaussian radial basis function. To determine these quantities, we propose a tailor-made version of generalized information criteria (GIC) constructed in the same manner as in Konishi and Kitagawa (1996). In the final section, we show worked examples with Japanese governmental bond data. By the use of bootstrapping, we demonstrate that the forward rate curve estimated by our method is much more stable than the one derived by the cubic spline model with cross validation. It is also shown that even in a small sample case our model with Gaussian basis still gives rise to a considerably stable forward rate compared to those obtained by natural cubic spline.