

実対称定値一般固有値問題に対する レゾルベントの多項式によるフィルタの構成法の検討

村上 弘^{1,a)}

概要: フィルタ対角化法を用いると、実対称定値一般固有値問題の指定した範囲に固有値がある固有対を求めることができる。本論文では、この方法に用いるフィルタ作用素として、ある虚数をシフトとする 1 個のレゾルベントの多項式を考察する

キーワード: フィルタ対角化, 固有値問題, レゾルベント, 多項式

Considerations on the Constructions of the Filter which is a Polynomial of Resolvent to Solve a Real Symmetric Definite Generalized Eigenproblem

HIROSHI MURAKAMI^{1,a)}

Abstract: For a real symmetric definite generalized eigenproblem, by the use of the filter diagonalization method, we can solve those eigenpairs whose values are in the specified interval. In this report, we consider the filter operator which is a polynomial of a resolvent whose shift is an imaginary number.

Keywords: filter diagonalization, eigenproblem, resolvent, polynomial

1. 概要

フィルタ対角化法を用いて、疎行列の一般固有値問題 $A\mathbf{x}=\lambda B\mathbf{x}$ の固有値が指定した範囲にある固有対を解くことができる。これまでは、シフトが異なるレゾルベントの線形結合をフィルタに用いる方法を研究してきた [1], [2].

用いるシフトは 6 個から 16 個程度である。シフトが τ のレゾルベント $\mathcal{R}(\tau) \equiv (A - \tau B)^{-1}B$ の作用は、 $A - \tau B$ を係数とする連立 1 次方程式を解いて実現する。連立 1 次方程式は、行列が帯なら LU 分解による直接法で解き、一般疎なら不完全 LU 分解を前処理に用いた反復法で解く。

すると全体の計算量の中で LU 分解の計算量がかなりの割合を占める。また分解された因子 L や U を連立 1 次方

程式を解き終わるまで保持するための記憶容量も、この方法を適用する上での制約要因となりがちである。それらに対処する簡単な方法として、

- 並行に処理するレゾルベントの個数を減らす。
- 使用するレゾルベントの個数を減らす。そうして得られる低次のフィルタを反復する。

などが容易に思いつくので、まずそれらについて以下に述べる。

● 並行に処理するレゾルベントの個数を減らす方法

複数のレゾルベントの適用を何個ずつかの組に分割して、各組を順次に（各組の内側では並行に）実行すれば、 LU 分解の保持に必要な記憶容量を減らせる。ただし分解にかかる演算量の総合計は変わらない。また経過時間は分割した数に比例して延びてしまう。この方法は計算の並行度の変更だけなので、得られる近似解の品質は変化しない。

¹ 首都大学東京・数理情報科学専攻
Department of Mathematics and Information Sciences,
Tokyo Metropolitan University
^{a)} mrkmhrsh@tmu.ac.jp

● 使用するレゾルベントの個数を減らす方法

フィルタを構成するレゾルベントの個数を減らすことで LU 分解の計算量を減らし、また分解された因子を保持するための記憶容量も減らす方法である。しかしそれにより得られる低次のフィルタの伝達特性は悪化する。特に、阻止域に於ける減衰率が不十分となるので、その低次のフィルタを数回反復したものを対角化用のフィルタとして用いる。低次のフィルタを最初に作用させる際の LU 分解を保持すれば、それ以降の反復では新たに分解を行う必要はない。

但し、レゾルベントの個数を極端に少なくすると、阻止域で減衰率が十分になるまで反復したフィルタの伝達特性は、通過域での伝達率の最大最小の比が大きくなり、得られる近似解の精度が悪化する。

今回、記憶容量が強い制約条件である場合を想定するが、それは上記のいずれの方針とも異なり、使用するレゾルベントは 1 個だけとする（シフトは虚数にする）。そうしてフィルタ作用素はそのレゾルベントの多項式を利用して構成する。そのためレゾルベントの作用を与える連立 1 次方程式を解くのに必要な LU 分解あるいは不完全 LU 分解も 1 回になる。フィルタ作用素がレゾルベントについて n 次多項式ならば、その作用を実現する過程で全部でレゾルベントの適用が n 回現れるが、最初の LU 分解の結果が保持できれば分解は 1 回で済む。

2. レゾルベントの多項式のフィルタ

いま解きたい実対称定値一般固有値問題を $A\mathbf{v} = \lambda B\mathbf{v}$ とする。行列 A と B は実対称、 B は正定値である。

シフト τ のレゾルベントを $\mathcal{R}(\tau) \equiv (A - \tau B)^{-1}B$ とすると、レゾルベント $\mathcal{R}(\tau)$ の固有対 $(\lambda, \mathbf{v})$ への作用は $\mathcal{R}(\tau)\mathbf{v} = 1/(\lambda - \tau) \cdot \mathbf{v}$ である。いまフィルタ $\mathcal{F}$ がこのレゾルベントの n 次多項式（の実部）（ c_∞ は実数）で以下のように与えられるとする：

$$\mathcal{F} = c_\infty I + \operatorname{Re} \sum_{k=1}^n \gamma_k \{\mathcal{R}(\tau)\}^k.$$

固有対 $(\lambda, \mathbf{v})$ に対するフィルタの作用は $\mathcal{F}\mathbf{v} = f(\lambda)\mathbf{v}$ となることがわかる。ここで $f(\lambda)$ は λ の実有理関数

$$f(\lambda) = c_\infty + \operatorname{Re} \sum_{k=1}^n \gamma_k / (\lambda - \tau)^k$$

であり、固有値 λ の固有ベクトルのフィルタによる伝達率（伝達関数）を与える。

いま区間 $[a, b]$ に固有値がある固有対を求めたいとする。そのとき $\lambda \in [a, b]$ から $t \in [-1, 1]$ への 1 次変換 $\lambda = (b+a)/2 + (b-a)t/2$ で λ の正規化座標 t を定義する。すると $t \in [-1, 1]$ は通過域、 $1 < \mu \leq |t|$ は阻止域、途中の

$1 < |t| < \mu$ は遷移域である。座標 t の伝達関数 $g(t)$ を $g(t) \equiv f(\lambda)$ で定義する。伝達関数 $g(t)$ の値分布への制約は、通過域での $g(t)$ の最小値は g_{pass} で、通過域にあるときに限り $g(t) \geq g_{\text{pass}}$ であり、阻止域に於ける $|g(t)|$ の上限値は g_{stop} である。さらに $g(t)$ を偶関数に限ると、極は純虚数の複素共役対である。共役対をなす極のうちで虚部が正のものを $\sigma\sqrt{-1}$ ($\sigma > 0$) とすると、

$$g(t) = c'_\infty + \operatorname{Re} \sum_{k=1}^n c_k / (t - \sigma\sqrt{-1})^k$$

である。もとの座標 λ での伝達関数は

$$f(\lambda) = c_\infty + \operatorname{Re} \sum_{k=1}^n \gamma_k / (\lambda - \tau)^k$$

であることと $f(\lambda) \equiv g(t)$ より両者の間には、定数項について $c_\infty = c'_\infty$ 、極の位置について $\tau = (a+b)/2 + \sigma(b-a)/2 \cdot \sqrt{-1}$ 、多重極の係数について $\gamma_k = c_k \{(b-a)/2\}^k$ 、 $k=1, 2, \dots, n$ の関係がある。すると $g(t)$ が決まれば $f(\lambda)$ が決まり、 $f(\lambda)$ に対応してフィルタ作用素 $\mathcal{F}$ も決まる。

さらに計算の簡便化のため、実数 $\sigma > 0$ による尺度変換で、 $t \equiv \sigma x$ と置き、 $h(x) \equiv g(t)$ と定義する：

$$h(x) = c''_\infty + \operatorname{Re} \sum_{k=1}^n \alpha_k / (1 + x\sqrt{-1})^k.$$

すると係数の間に $c''_\infty = c_\infty$ および $c_k = (-\sigma\sqrt{-1})^k \alpha_k$ 、 $k=1, 2, \dots, n$ の関係がある。また $g(t)$ が偶関数であることから $h(x)$ も偶関数であり、係数 α_k はすべて実数であることがわかる。さらに $x \equiv \tan \theta$ 、 $\hat{h}(\theta) \equiv h(x)$ とおくと、

$$\hat{h}(\theta) = c_\infty + \sum_{k=1}^n \alpha_k \cos(k\theta) (\cos \theta)^k.$$

である。以降では無限遠での伝達率を零に限定して $c_\infty = 0$ とする。

ここまですべて整理すれば（ただし $\sigma = 1$ と簡略化）。

- 1 次変換 $\lambda = (a+b)/2 + x(b-a)/2$ で $\lambda \in [a, b]$ を $x \in [-1, 1]$ に対応させる。
- 実有理関数 $h(x)$ が x の偶関数で、 $x = \pm\sqrt{-1}$ だけに極を持ち、（簡単のため）無限遠での値を零とすると、

$$h(x) = \operatorname{Re} \sum_{k=1}^n \alpha_k / (1 + x\sqrt{-1})^k.$$

- 通過域 $[-1, 1]$ 付近から離れると $|h(x)|$ が急減少するように、実パラメタ α_k 、 $k=1, 2, \dots, n$ を調整する。
- $h(x)$ の座標を λ に変換した実有理関数を $f(\lambda) \equiv h(x)$ とすると、伝達関数が $f(\lambda)$ のフィルタ作用素 $\mathcal{F}$ はレゾルベントの多項式を用いて

$$\mathcal{F} = \operatorname{Re} \sum_{k=1}^n \gamma_k \{\mathcal{R}(\tau)\}^k$$

である。ただしシフトは $\tau \equiv (a+b)/2 + (b-a)/2 \cdot \sqrt{-1}$,
係数は $\gamma_k \equiv \alpha_k \{(b-a)/2\}^k (-\sqrt{-1})^k$ である。

次数 n を与え、極が $\pm\sqrt{-1}$ だけにある偶関数の n 次実有理関数

$$h(x) = \text{Re} \sum_{k=1}^n \alpha_k / (1 + x\sqrt{-1})^k$$

を ($h(x)$ が偶関数であることから α_k は実数), 実数パラメタ α_k , $k=1, 2, \dots, n$ をうまく調整することで, $|h(x)|$ の値が通過域 $[-1, 1]$ では 1 の付近にあり, 通過域からある程度離れた阻止域では微小となるようにする。計算の便宜上, $x \equiv \tan \theta$ と置換して, $\hat{h}(\theta) \equiv h(x)$ と置くと

$$\hat{h}(\theta) = \sum_{k=1}^n \alpha_k \cos(k\theta) (\cos \theta)^k.$$

偶関数性から $\theta \geq 0$ だけを考えれば良いから, 通過域も $\theta \in [0, \pi/4]$ だけを考えれば良い。

以下では, 実数パラメタ α_k の調整を最小 2 乗法 (の系統) で行う場合を示す。

2.1 最小 2 乗法 (解法 I)

阻止域と通過域に於ける伝達関数 $\hat{h}(\theta)$ の理想特性から誤差の 2 乗を (簡単のため重み 1 で) θ で積分した値をそれぞれ J_{stop} と J_{pass} とする。そうして微小な正数 η を適切に選び, 目的関数 $J \equiv J_{\text{stop}} + \eta J_{\text{pass}}$ の最小化を行なう。

通過域の中央と端点 $x=0$, $x=1$, 阻止域の両端点 $x=\mu$, $x=\infty$ に対応する θ の値をそれぞれ $\theta_0=0$, $\theta_1=\pi/4$, $\theta_\mu=\arctan \mu$, $\theta_\infty=\pi/2$ とすると,

$$J_{\text{stop}} \equiv \int_{\theta_\mu}^{\theta_\infty} \{\hat{h}(\theta)\}^2 d\theta = \sum_{p,q=1}^n \alpha_p A_{p,q} \alpha_q$$

および

$$J_{\text{pass}} \equiv \int_{\theta_0}^{\theta_1} \{1 - \hat{h}(\theta)\}^2 d\theta = \sum_{p,q=1}^n \alpha_p B_{p,q} \alpha_q - 2 \sum_{p=1}^n \alpha_p B_{p,0} + \text{Const}$$

である。ただしここで

$$A_{p,q} \equiv \int_{\theta_\mu}^{\theta_\infty} \cos(p\theta) \cos(q\theta) (\cos \theta)^{p+q} d\theta,$$

$$B_{p,q} \equiv \int_{\theta_0}^{\theta_1} \cos(p\theta) \cos(q\theta) (\cos \theta)^{p+q} d\theta$$

である。これら $A_{p,q}$, $B_{p,q}$ を与える定積分の値は解析的な表式を用いて計算できる。そうして $J \equiv J_{\text{stop}} + \eta J_{\text{pass}}$ の最小化条件は, $b_p = B_{p,0}$ とおくと, 係数行列が実対称正定値の連立 1 次方程式

$$(A + \eta B) \alpha = \eta b$$

になるので, これを解けば α_k , $k=1, 2, \dots, n$ が求まる。

2.2 最小 2 乗法 (解法 II)

α を 2-ノルム一定のベクトルとみなして, まず先に阻止域での積分

$$J_{\text{stop}} = \sum_{p,q=1}^n \alpha_p A_{p,q} \alpha_q = \alpha^T A \alpha$$

の最小化を試みる。実対称正定値行列 A の固有分解を $A \rightarrow U D U^T$ とする。ただし D は A の固有値を並べた対角行列で, U は固有値の順に固有ベクトルを並べた直交行列である。 J_{stop} を最小化するのは固有値が最小の固有ベクトル α である。しかしこれでは通過域での近似程度を調整する自由度が無いので, 調整可能な自由度を含められるように変更を加える。

いま微小な正数 ϵ を適切に決める。固有値が ϵ 以下の A の固有対が ℓ 個存在したとする。そのとき $\mathcal{S}^{(\ell)}$ をそれら ℓ 個の固有ベクトルで張られた部分空間とする。部分空間 $\mathcal{S}^{(\ell)}$ 内にベクトル α があれば, $J_{\text{stop}} \leq \epsilon \|\alpha\|_2^2$ である。そこで α を部分空間 $\mathcal{S}^{(\ell)}$ 内に制限して目的関数 J_{pass} の値を最小化する。

これは ℓ 次に縮小された実対称正定値の連立 1 次方程式に帰着できる。通過域での特性の誤差を減らすべく部分空間を捻げるとその代りに阻止域での特性が悪化 (伝達率の大きさが増大) する。

2.3 実験例

解法 II により $n=15$, $\mu=2.0$ のフィルタ係数を求めた例を示す。最小化計算の過程には IEEE 754 規格の 128 ビット四倍精度浮動小数点数とその演算を使用した。行列 A の固有分解には threshold Jacobi 法を使用した。この例では固有値の切断閾値を $\epsilon=10^{-30}$ と設定すると $\ell=2$ となった。得られた係数 α_k , $k=1, 2, \dots, n$ を倍精度に丸めた値を表 1 に示す。係数の大きさは 1 程度で良く揃っている。偶関数性から, 片側 $x>0$ での伝達関数

$$h(x) = \text{Re} \sum_{k=1}^n \alpha_k / (1 + x\sqrt{-1})^k$$

の大きさ $|h(x)|$ を (係数 α_k を IEEE 754 の 64bit 倍精度に丸めてから倍精度演算を行ない) 両対数でプロットしたグラフを図 1 に示す。

伝達率 $h(x)$ は最大値 1.325 を $x=0.216$ 付近でとり, $h(0)=7.2 \times 10^{-1}$, $h(1)=2.3 \times 10^{-4}$, $h(1.5)=1.9 \times 10^{-8}$, $h(2)=1.1 \times 10^{-15}$ となった。通過域での $h(x)$ の最小値は $y_{\min}=0.71862$ で, 遷移域で $h(x)=y_{\min}$ となる x の値は $x_0=0.388220$ として, 再スケール処理 $h'(x') \equiv h(x' \cdot x_0)/y_{\min}$ を行なうと, $x' \in [-1, 1]$ では $h'(x') \geq 1$ で, 最大値 1.844 を $x'=0.557$ でとり, $h'(1)=1$, $h'(2)=1.18 \times 10^{-2}$, $h'(3)=1.84 \times 10^{-5}$, $h'(4)=8.68 \times 10^{-9}$, $h'(5)=1.18 \times 10^{-13}$, $h'(6)=2.94 \times 10^{-16}$ など。そうして $\mu=3$ にすると

表 1 係数 α (解法 II, $n=15$, $\mu=2.0$, $\epsilon=10^{-30}$)

k	α_k
1	3.10422 91727 23495 E-1
2	3.10422 91727 25609 E-1
3	2.85453 67519 83506 E-1
4	2.35515 19113 67395 E-1
5	1.64913 99494 59607 E-1
6	8.22631 58940 55446 E-2
7	-6.57520 79352 44120 E-4
8	-7.11802 27019 60262 E-2
9	-1.18756 19212 14338 E-1
10	-1.37828 28527 33139 E-1
11	-1.29654 88587 73316 E-1
12	-1.01680 66293 50991 E-1
13	-6.60360 83956 00963 E-2
14	-3.26587 11429 62141 E-2
15	-1.19174 53737 97113 E-2

$g_{\text{stop}}=1.8 \times 10^{-5}$, $\mu=4$ にすると $g_{\text{stop}}=8.7 \times 10^{-9}$, $\mu=5$ にすると $g_{\text{stop}}=1.2 \times 10^{-13}$, $\mu=6$ にすると $g_{\text{stop}}=2.9 \times 10^{-16}$ などとなる.

μ を大きくして遷移域が広がると, 不要な固有値が遷移域にそれだけ多く入りやすくなり, フィルタで濾過するベクトルの個数を増やす必要性が高まる. また g_{stop} が微小でなければ, 不変部分空間の近似度は低下する. 通過域内での伝達率の最大最小比が大きいと, 計算で得られた固有値が通過域内にある近似固有対の精度の均一性が低下する可能性がある.

3. ここまでのまとめ

「フィルタ対角化法」は, レゾルベントから構成されたフィルタを用いて, 不変部分空間の基底の近似を構成する. レゾルベントの作用は, 連立 1 次方程式を直接法もしくは反復法で解いて実現する. 直接法なら LU 分解, 反復法も前処理に不完全 LU 分解を行えば, 異なるシフトのレゾルベントそれぞれについて, LU 分解の計算と分解結果を保持する記憶が必要になる. 高品位のフィルタの構成にはある程度の個数のレゾルベントが必要になるが, すべてのレゾルベントを一度に作用させると, LU 分解の保持にレゾルベントの個数分だけの記憶容量が必要になって実行の制約となりがちである. すると全てのレゾルベントの作用を一斉に並行して計算するのではなくて, 何組かに分割して各組ごとに順次に計算をする必要が出てくる.

ここまでは, レゾルベントを最小個数の 1 個として, レゾルベントの多項式によりフィルタを構成する可能性を調べた. そうして伝達関数が理想特性を良く近似するように, 最小 2 乗法の手法でレゾルベントの多項式の係数を決めた. それにより得られるフィルタの伝達関数は, 自由に極を配置できる場合に比べると, 遷移域で急峻な値の変化ができない. そのうちに構成したフィルタを用いてフィル

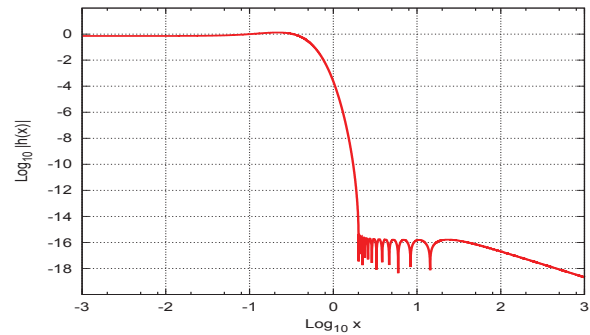


図 1 伝達関数の大きさ $|h(x)|$ (解法 II, $n=15$, $\mu=2.0$, $\epsilon=10^{-30}$. (両軸対数))

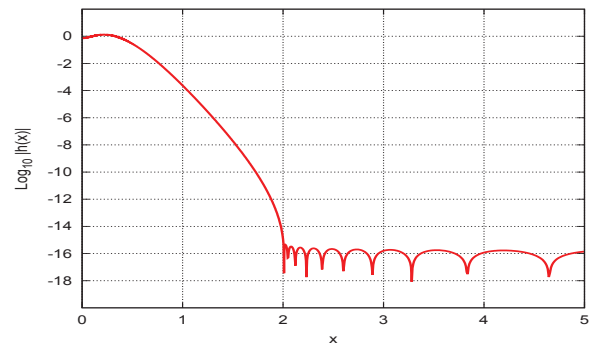


図 2 伝達関数の大きさ $|h(x)|$ (解法 II, $n=15$, $\mu=2.0$, $\epsilon=10^{-30}$. (縦軸対数))

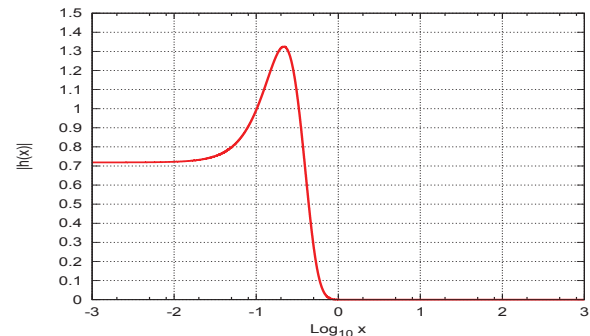


図 3 伝達関数の大きさ $|h(x)|$ (解法 II, $n=15$, $\mu=2.0$, $\epsilon=10^{-30}$. (横軸対数))

タ対角化法を実験をして, 近似固有対がうまく得られるか検証を行なうつもりである.

4. 伝達関数を明示的な数式で直接与える試み

ここまでは, 数値最適化法を用いて理想特性からのフィルタの伝達関数の誤差のノルムを最小化して伝達関数の係数を決める手法であった. 最適化を用いる方法には汎用性があるが, 数値的悪条件による困難さや見通しが良くないのが難点である. そこで今回は以下で, 最適化法に頼らず明示的な数式で理想特性を近似する伝達関数を直接与えることを試みる.

いま固有値の座標を 1 次変換した正規化座標を t とする. シフトが $t=\sqrt{-1}$ のレゾルベントの伝達関数は $1/(t-\sqrt{-1})$ で, $\text{Im } 1/(t-\sqrt{-1}) = 1/(t^2+1)$ であるから, 「実ベクトル

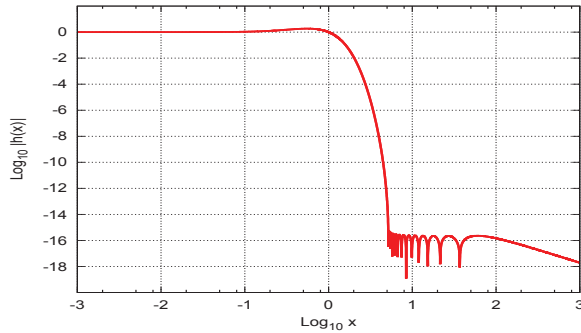


図 4 伝達関数の大きさ $|h'(x)|$ (解法 II, $n=15$, 再スケール. (両軸対数))

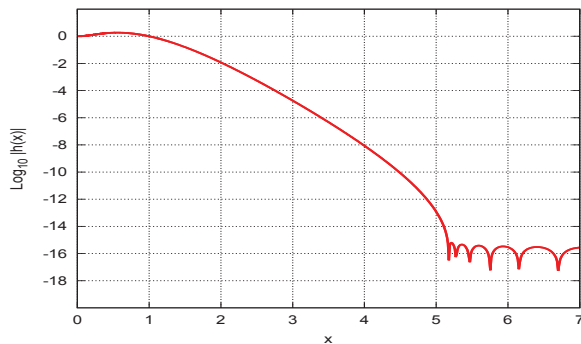


図 5 伝達関数の大きさ $|h'(x)|$ (解法 II, $n=15$, 再スケール. (縦軸対数))

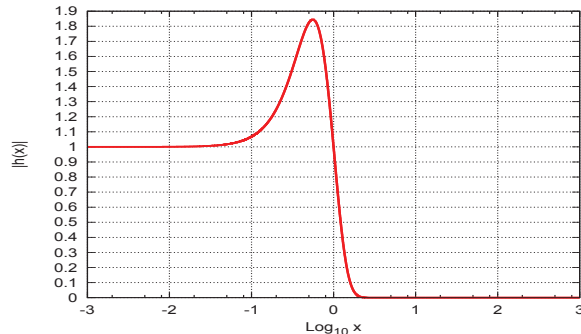


図 6 伝達関数の大きさ $|h'(x)|$ (解法 II, $n=15$, 再スケール. (横軸対数))

にレゾルベントを作用させて虚部をとる」操作を Y とするとその作用の伝達関数は $1/(t^2+1)$ である。そこでいま $y=1/(t^2+1)$ とおき、フィルタの伝達関数を y の多項式とする。 t が実数ならば $y \in (0, 1]$ である。 $t=0$ で $y=1$ であり、 $t=1$ で $y=1/2$ であり、 $t=\infty$ で $y=0$ である。すると理想フィルタの伝達関数の阻止域は $y \in [0, 1/2)$ に、通過域は $y \in (1/2, 1]$ に対応する (図 7)。

4.1 規格化 beta 関数による構成

不完全 beta 関数 $B(p, q; y)$ は p, q を正の実数とするとき、

$$B(p, q; y) \equiv \int_0^y z^{p-1} (1-z)^{q-1} dz$$

で定義される [3]。また (完全) beta 関数は $B(p, q) \equiv B(p, q; 1) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$ である。そうして規格化

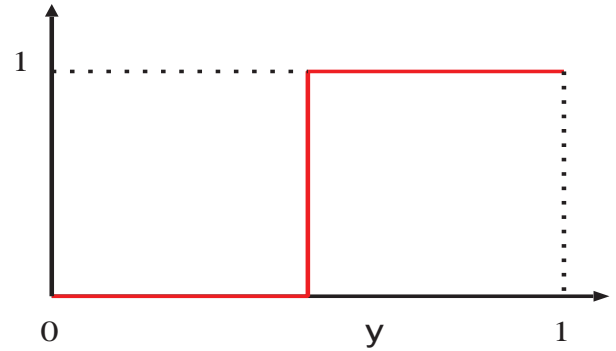


図 7 理想フィルタの伝達特性 (引数 y)

された不完全 beta 関数 (以下では「規格化 beta 関数」と呼ぶ) を

$$I(p, q; y) \equiv B(p, q; y)/B(p, q)$$

と定義する。 $I(p, q; y)$ は $y \in [0, 1]$ で単調増加で、 $y=0$ で値は 0, $y=1$ で値は 1 である。また $y=0$ の近傍で $O(y^p)$, $y=1$ の近傍で $1-O((1-y)^q)$ である。 $I(p, q; y) + I(q, p; 1-y) = 1$ であることもわかる。特に $p=q$ の場合は $y=1/2$ に於いて $I(p, p; 1/2) = 1/2$ である。特に p, q が正の整数の場合には、規格化 beta 関数 $I(p, q; y)$ は次数 $n=p+q-1$ の多項式で、二項係数を用いて以下の式で表せる。

$$I(p, q; y) = \sum_{j=p}^n \binom{n}{j} y^j (1-y)^{n-j}$$

$y=1/(t^2+1)$ に対応する作用 Y と、 $1-y$ に対応する作用 $I-Y$ は、同じシフトのレゾルベントの作用 (と虚部をとる操作) を用いて実現できる。 y^l や $(1-y)^m$ に対応する作用も繰り返しにより実現できる。 y の中で展開すると高次では係数膨張により精度低下を招くので、 $y^j (1-y)^{n-j}$ の項のままで展開しないで用いる方がよい。

伝達関数が $I(p, q; y)$ であるフィルタはいわゆるバターワース特性を持ち、 p が大きいほど無限遠 $t = \pm\infty$ 付近での阻止性能が良く、 q が大きいほど原点 $t = 0$ 付近での平坦性が良い。

4.1.1 規格化 beta 関数の帰納的関係の応用

表式

$$I(p, q; Y) = \sum_{j=p}^n \binom{n}{j} Y^j (I-Y)^{n-j}$$

を直接用いてベクトルへの作用を計算すると、 Y の指数 j を増やすとき $(I-Y)$ の指数 $n-j$ は減るので、作業用のベクトルが多く必要になる。そこで $I(p, q; y)$ ($p \geq 1, q \geq 1$) の帰納的な関係 $I(1, 1; y) = y$, $I(p+1, q; y) = I(p, q; y) - \binom{p+q-1}{p} y^p (1-y)^q$ および $I(p, q+1; y) = I(p, q; y) + \binom{p+q-1}{q} y^p (1-y)^q$ を用いる。例えば

$$I(1, 1; Y) = Y,$$

$$I(2, 1; Y) = I(1, 1; Y) - \left(\frac{1}{1}\right)Y(I-Y), \quad ! (I-Y) \text{ を乗じる.}$$

$$I(2, 2; Y) = I(2, 1; Y) + \left(\frac{2}{1}\right)Y^2(I-Y), \quad ! Y \text{ を乗じる.}$$

$$I(3, 2; Y) = I(2, 2; Y) - \left(\frac{3}{2}\right)Y^2(I-Y)^2, \quad ! (I-Y) \text{ を乗じる.}$$

$$I(3, 3; Y) = I(3, 2; Y) + \left(\frac{4}{2}\right)Y^3(I-Y)^2, \quad ! Y \text{ を乗じる.}$$

$$I(4, 3; Y) = I(3, 3; Y) - \left(\frac{5}{3}\right)Y^3(I-Y)^3, \quad ! (1-Y) \text{ を乗じる.}$$

$$I(4, 4; Y) = I(4, 3; Y) + \left(\frac{6}{3}\right)Y^4(I-Y)^3, \quad ! Y \text{ を乗じる.}$$

のように、2つの正整数 p, q を単調に増加させる経路に対応して、ベクトルに Y あるいは $I-Y$ の作用を積み重ねるように適用しながら二項係数を乗じて加えていくと、多項式 $I(p, q; y)$ を伝達特性に持つフィルタが実現できて、中間作業用のベクトルも2個でよい。この再帰的關係を用いた計算の数値安定性や、 p, q の経路をどのように選ぶと数値的性質が良いかについては今後調査する必要がある。

例 その1

例として、次数 $n=29$ の規格化 beta 関数による (引数を t とする) 伝達関数を求めた。それに対して尺度変換を用いて $t=1$ に於ける伝達率を $1/2$ に調整する前の伝達率のグラフ (図 8, 図 9, 図 10, 図 11) と調整した後の伝達率のグラフ (図 12, 図 13, 図 14, 図 15) をそれぞれ示す。

例 その2

規格化 beta 関数による伝達関数の例を示す。但し、 $n=39$ で、これも尺度変換により $t=1$ で伝達率が $1/2$ になるように調整する前 (図 16, 図 17, 図 18, 図 19) と調整した後 (図 20, 図 21, 図 22, 図 23) のグラフをそれぞれ示す。

4.1.2 規格化 beta 関数の別の表現法の考察

いま $x \in [-1, 1]$ とする (変換 $x=2y-1$ で $y=1/(t^2+1) \in [0, 1]$ に対応)。簡単のため $p=q=m+1$ に限定し、 $I(p, p; y = (x+1)/2)$ と一致する多項式 $P(x)$ を

$$P(x) \equiv 1/2 + x \sum_{j=0}^m a_j (1-x^2)^j$$

とおくと、 $P(0)=1/2$ であり、 $P(x)-1/2$ は奇関数で、その次数は $2m+1$ である。 $a_0=1/2$ とすると $P(-1)=0, P(1)=1$ 。

両端点 $x=\pm 1$ での m 階以下の導関数の値 $P^{(j)}(\pm 1)$, $j=1, 2, \dots, m$ が零となるように係数 $a_m, m=1, 2, \dots$ を順番に決定できて、 $a_1 = 1/4, a_2 = 3/16, a_3 = 5/32, a_4 = 35/256, a_5 = 63/512, a_6 = 231/2048, a_7 = 429/4096, a_8 = 6435/65536, a_9 = 12155/131072$, などとなる。係数 a_m は非常に緩やかに減少する (表 2, 図 24)。この係数を決める計算には数式処理システム Maxima を用いた。

図 25, 図 26 に、この方法で得られた各 m の値に対する伝達関数のグラフを横軸を t としてプロットした ($x=2y-1, y=1/(t^2+1)$)。

4.2 不完全 beta 関数の模倣による拡張の試み

不完全 beta 関数は、 $[0, 1]$ 区間内での階段関数の近似で

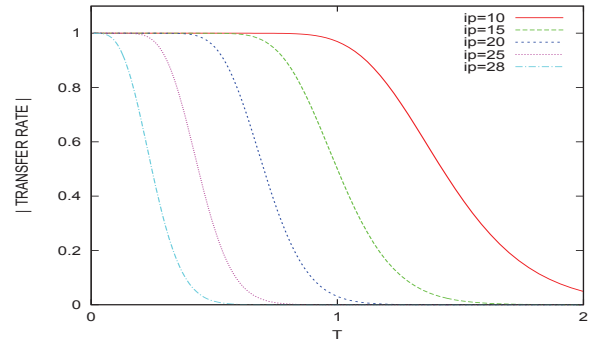


図 8 規格化 beta 関数 (次数 $n=29$, 引数 t , 調整前) (両軸真数)

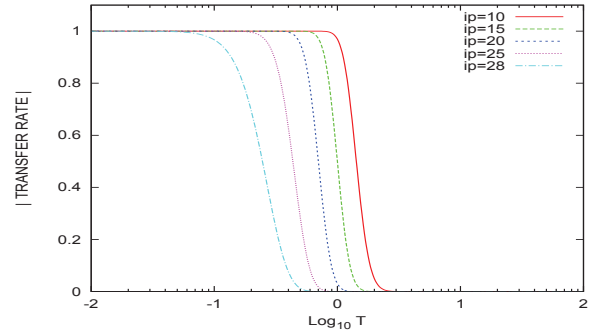


図 9 規格化 beta 関数 (次数 $n=29$, 引数 t , 調整前) (横軸対数)

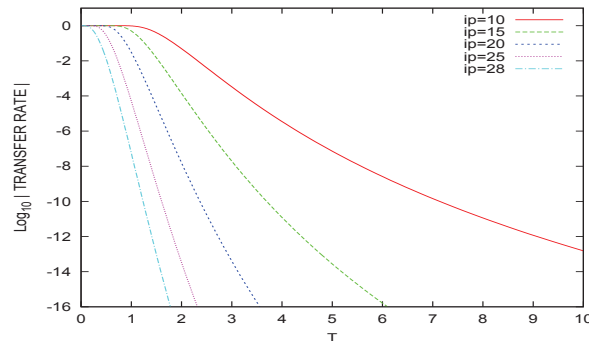


図 10 規格化 beta 関数 (次数 $n=29$, 引数 t , 調整前) (縦軸対数)

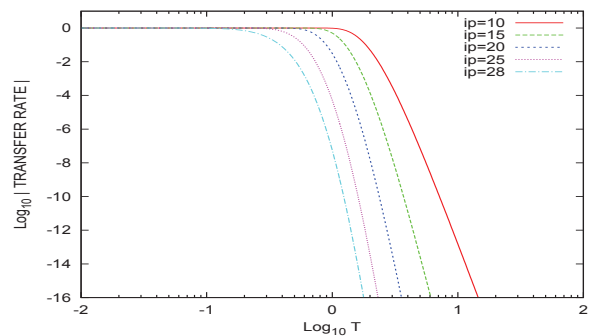


図 11 規格化 beta 関数 (次数 $n=29$, 引数 t , 調整前) (両軸対数)

あり、その微分 $x^p(1-x)^q$ はデルタ関数の近似とみなせる。微分 $x^p(1-x)^q$ は、 $x=0$ 近傍で値の小さい因子 x^p と $x=1$ 近傍で値の小さい因子 $(1-x)^q$ の積であり、区間中央の $x=p/(p+q)$ 付近だけで大きな値をとる (図 27)。この構成を模倣する拡張を考える。

いま $x=0$ の近くで値の小さい因子として、 $[0, (1-\nu)/2]$ を基本区間とする p 次の直交多項式 $\Phi_p(x/c-1)$ を採用す

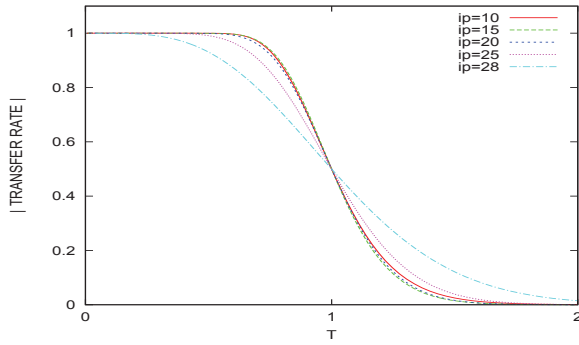


図 12 規格化 beta 関数 (次数 $n=29$, 引数 t , 調整後) (両軸真数)

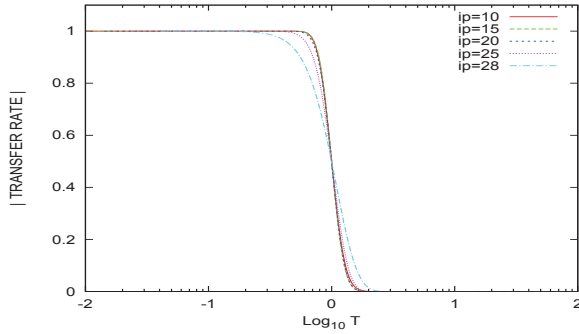


図 13 規格化 beta 関数 (次数 $n=29$, 引数 t , 調整後) (横軸対数)

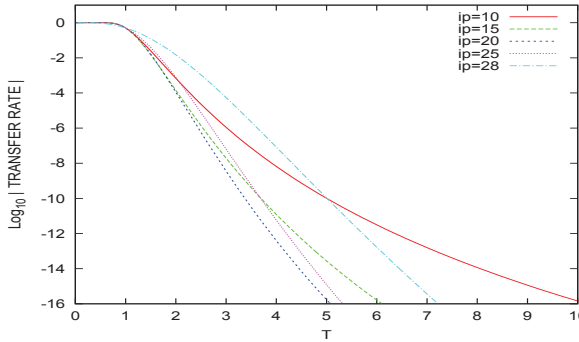


図 14 規格化 beta 関数 (次数 $n=29$, 引数 t , 調整後) (縦軸対数)

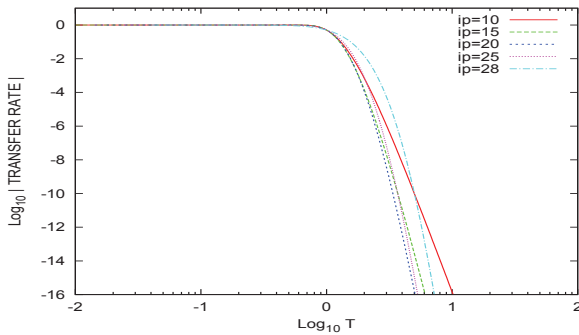


図 15 規格化 beta 関数 (次数 $n=29$, 引数 t , 調整後) (両軸対数)

る。ただし $c=(1-\nu)/4$ であり、 Φ_p の引数は基本区間の左端では -1 ，右端では 1 である。同様に、 $x=1$ の近くで値の小さい因子として、 $[(1+\nu)/2, 1]$ を基本区間とする q 次の直交多項式 $\Phi_q((1-x)/c-1)$ を採用する。 Φ_p の引数は基本区間の左端では 1 ，右端では -1 である。区間 $[0, 1]$ 内の両因子の基本区間の外部の区間 $((1-\nu)/2, (1+\nu)/2)$ でだけ、両因子の積 $\Phi_p(x/c-1) \cdot \Phi_q((1-x)/c-1)$ の値は比較

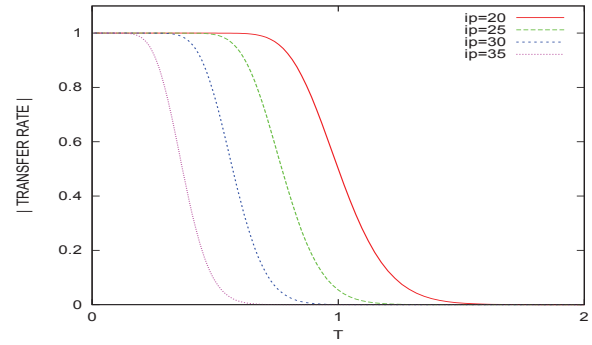


図 16 規格化 beta 関数 (次数 $n=39$, 引数 t , 調整前) (両軸真数)

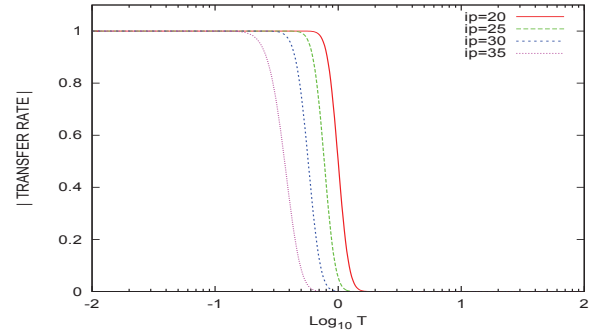


図 17 規格化 beta 関数 (次数 $n=39$, 引数 t , 調整前) (横軸対数)

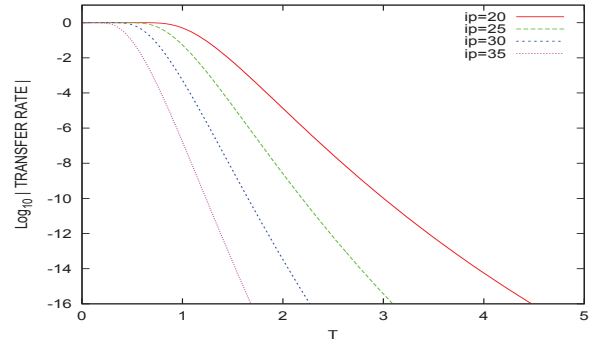


図 18 規格化 beta 関数 (次数 $n=39$, 引数 t , 調整前) (縦軸対数)

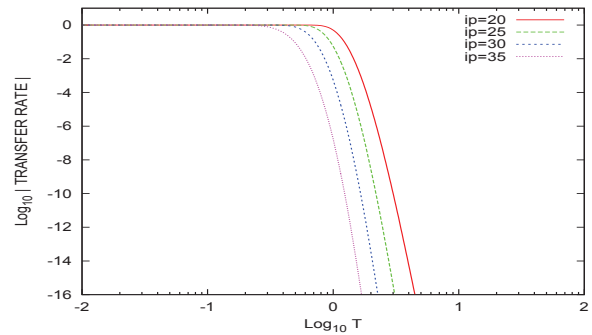


図 19 規格化 beta 関数 (次数 $n=39$, 引数 t , 調整前) (両軸対数)

的大で正となる (図 28)。

そこで、不完全 beta 関数の模倣により拡張された定義を

$$\tilde{B}(p, q; x) \equiv \int_0^x \Phi_p(z/c-1) \Phi_q((1-z)/c-1) dz$$

とする。それを規格化した関数を

$$\tilde{I}(p, q; x) \equiv \tilde{B}(p, q; x) / \tilde{B}(p, q; 1)$$

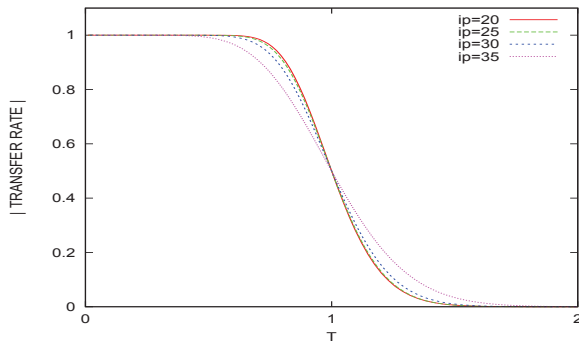


図 20 規格化 beta 関数 (次数 $n=39$, 引数 t , 調整後) (両軸真数)

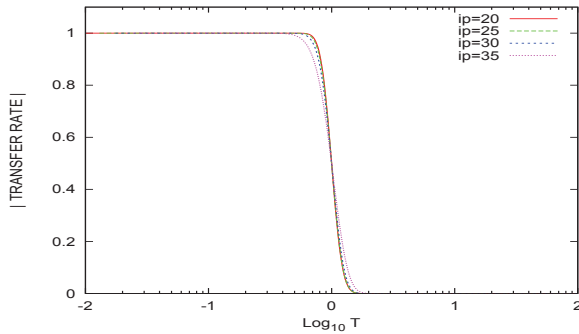


図 21 規格化 beta 関数 (次数 $n=39$, 引数 t , 調整後) (横軸対数)

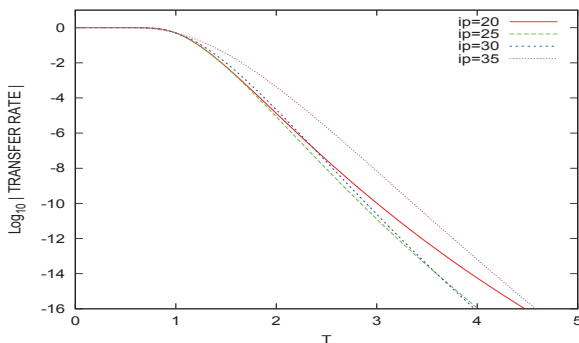


図 22 規格化 beta 関数 (次数 $n=39$, 引数 t , 調整後) (縦軸対数)

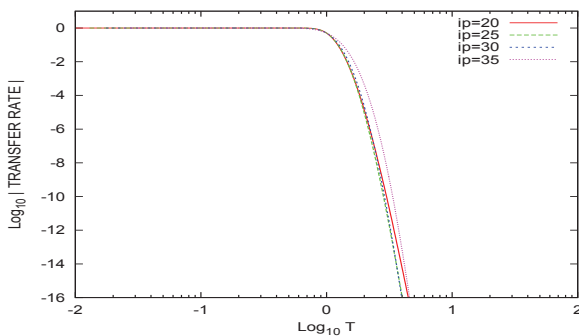


図 23 規格化 beta 関数 (次数 $n=39$, 引数 t , 調整後) (両軸対数)

とすると, $\tilde{I}(p, q; 0) = 0$, $\tilde{I}(p, q; 1) = 1$, $\tilde{I}(p, q; x) + \tilde{I}(q, p; 1-x) = 1$ である. 特に $p=q$ のときは $\tilde{I}(p, p; 1/2) = 1/2$ となる.

この拡張された方法を実施する場合の困難は, 多項式 $\tilde{I}(p, q; x)$ をたとえば x で巾展開すると, 高次では係数膨張により丸め誤差が拡大して精度が失われ, 結果が破綻することである. 実用化には多項式と対応する作用を数値的に

表 2 係数 a_m

m	a_m
0	0.5000000000000000
1	0.2500000000000000
2	0.1875000000000000
3	0.1562500000000000
4	0.1367187500000000
5	0.1230468750000000
6	0.1127929687500000
7	0.1047363281250000
8	0.0981903076171875
9	0.09273529052734375

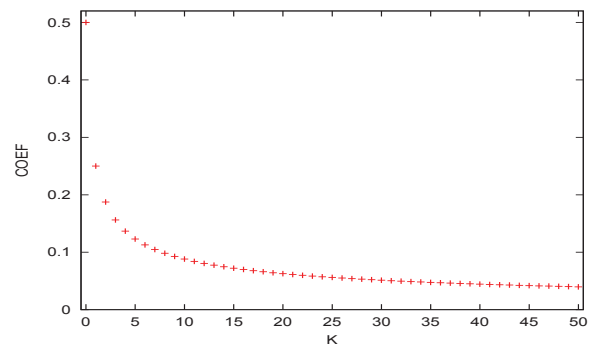


図 24 展開係数 a_m のグラフ

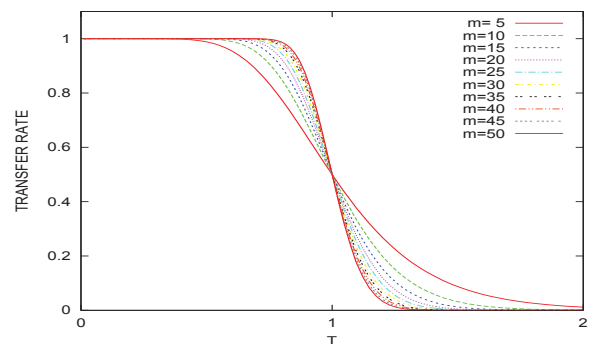


図 25 伝達関数のグラフ (引数は t , 両軸真数)

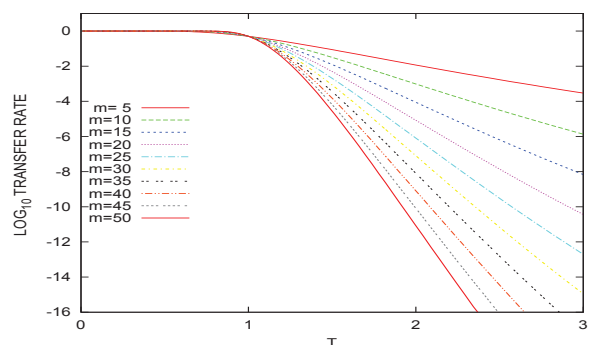


図 26 伝達関数のグラフ (引数は t , 縦軸対数)

安定に計算できる方法の開発が必要である.

例 その 1

図 29, 図 30, 図 31 に, 規格化 beta 関数を「拡張」した伝達関数 ($y=1/(t^2+1)$ の多項式) で $p=q=15$ の場合の

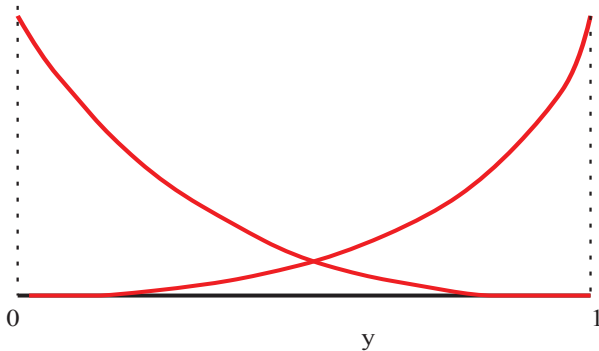


図 27 概念図. 不完全 beta 関数. 微分の各因子 (引数は y)

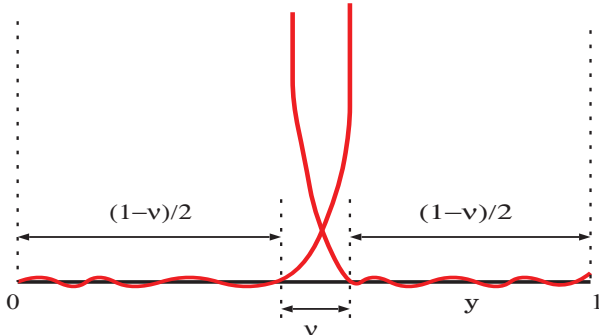


図 28 概念図. 不完全 beta 関数の「拡張」. 微分の各因子 (引数は y)

グラフを, 横軸に t の値をとりプロットしたものを示す.
「直交多項式」には第 2 種 Chebyshev 多項式を使用した.

例 その 2

図 32, 図 33, 図 34 に, 規格化 beta 関数を「拡張」した伝達関数 ($y=1/(t^2+1)$ の多項式) で $p=q=20$ の場合のグラフを, 横軸に t の値をとりプロットしたものを示す.
直交多項式には第 2 種 Chebyshev 多項式を使用した.

例 その 3

図 35, 図 36 に, 規格化 beta 関数を「拡張」した伝達関数 ($y=1/(t^2+1)$ の多項式) で $p=15, q=5$ の場合のグラフを, 横軸に t の値をとりプロットしたものを示す. 直交多項式には第 2 種 Chebyshev 多項式を使用した. (注: $p \neq q$ のとき 2 つの基本区間を同一幅にとるのはあまり良くない.)

4.2.1 不完全 beta 関数の「類似拡張」の別法の試み

これは $p=q$ の場合に該当するが, いま $x \in [-1, 1]$ として

$$\tilde{P}(x) = (1+x)/2 + x \sum_{j=1}^m a_j (1-x^2)^j$$

とすると, $\tilde{P}(0)-1/2$ は奇関数で, $\tilde{P}(-1)=0, \tilde{P}(1)=1$ である. 両端点の付近で直交多項式的な振る舞いを導入するために, 値が $x=-1$ の付近の選点集合上 (たとえば直交多項式の零点分布上) で $\tilde{P}(x)$ が零となる条件を課して, 展開係数 $a_j, j=1, 2, \dots$ を決めてみた.

図 37, 図 38 のグラフで, $m=10$ の場合 (赤線) には二つの基本区間 $[-1, 0.8], [0.8, 1]$ にそれぞれ 10 個の選点を

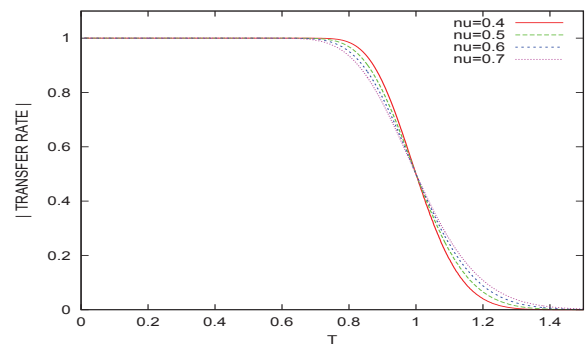


図 29 規格化 beta 関数の「拡張」($p=q=15$) (両軸真数)

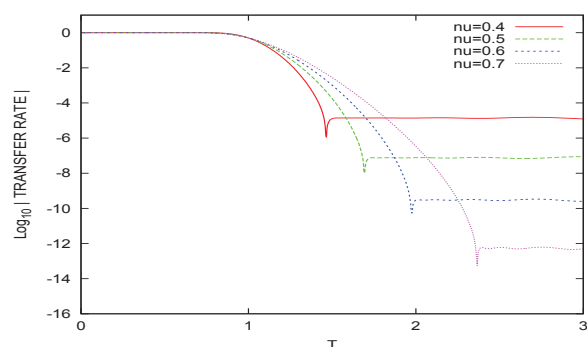


図 30 規格化 beta 関数の「拡張」($p=q=15$) (縦軸対数)

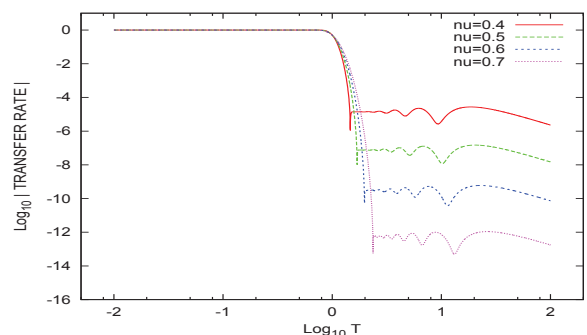


図 31 規格化 beta 関数の「拡張」($p=q=15$) (両軸対数)

割り当て, また $m=15$ の場合 (緑線) には二つの基本区間 $[-1, 0.6], [0.6, 1]$ にそれぞれ 15 個の選点を割り当てている. どちらも $t=1$ での伝達率は 0.5 で, $m=10$ の場合には $\mu=3$ 程度で $g_{\text{stop}}=10^{-8}$ 程度, $m=15$ の場合には $\mu=2$ 程度で $g_{\text{stop}}=10^{-8}$ 程度, という結果を得た. $y=1/(t^2+1)$ の特性を持つ作用 Y は, 実数ベクトルに対してレゾルベント $\mathcal{R}$ を作用させてから虚部をとることで実現する. $x=2y-1$ に対応する作用 X は, $2Y-I$ で実現できる. $(1-x^2)$ と対応する作用は, $I-X^2$ あるいは $(I-X)(I+X)$ として計算できる.

5. おわりに

フィルタ対角化法を用いて実対称定値一般固有値問題の指定された区間に固有値がある固有対を求めるのに必要な記憶量と演算量を低減することを目的として, シフトが虚数のレゾルベント $\mathcal{R}$ を一つだけ用いて, 「ベクトルに $\mathcal{R}$ を

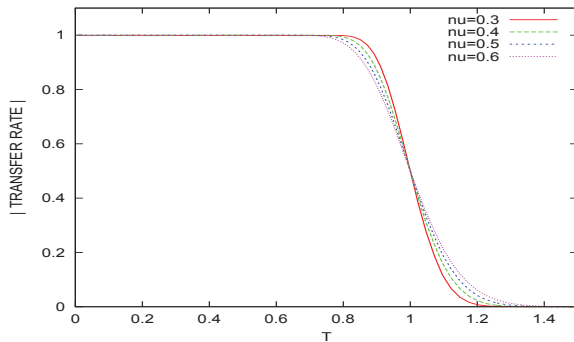


図 32 規格化 beta 関数の「拡張」($p=q=20$) (両軸真数)

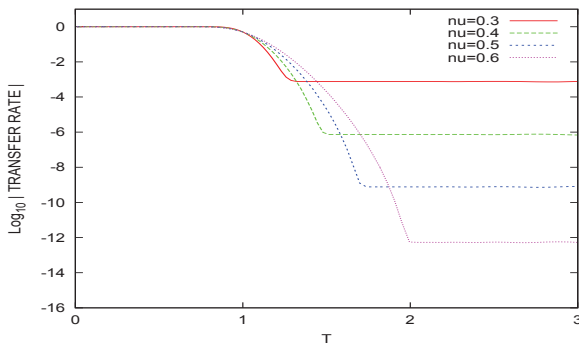


図 33 規格化 beta 関数の「拡張」($p=q=20$) (縦軸対数)

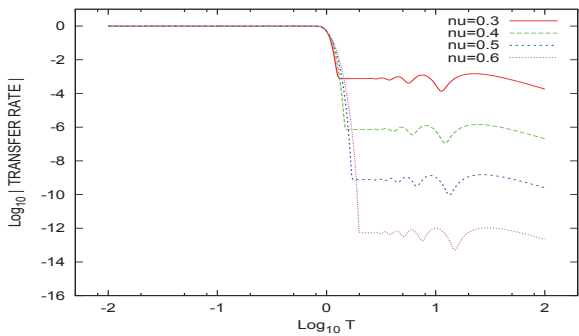


図 34 規格化 beta 関数の「拡張」($p=q=20$) (両軸対数)

作用させたのちに虚部をとる作用」の多項式によりフィルタ作用素を構成する方法について、今回新たに考察した。

「ベクトルに $\mathcal{R}$ を作用させたのちに虚部をとる作用」の伝達関数は固有値の正規化座標を t として $y=1/(t^2+1)$ である。理想フィルタの伝達関数を y の多項式で近似する手法として、近似誤差のノルムの最小化で未定係数を決める以前の方法から、今回新たに明示的な式で近似多項式を与える方法を試みた。

まだ準備や検討の段階であり、今後とも実証実験も含めて考察を重ねる必要がある。

参考文献

- [1] 村上 弘: 固有値が指定された区間内にある固有対を解くための対称固有値問題用のフィルタの設計, 情報処理学会論文誌: コンピューティングシステム (ACS31), Vol.3, No.3 (2010), pp.1-21.
- [2] 村上 弘: 対称一般固有値問題のフィルタ作用素を用いた不変部分空間の近似構成, 情報処理学会論文誌: コンピュー

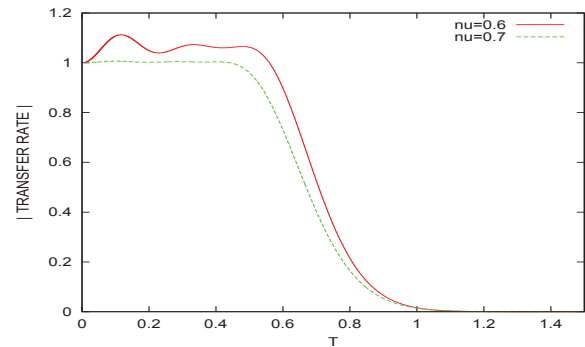


図 35 規格化 beta 関数の「拡張」($p=15, q=5$) (両軸真数)

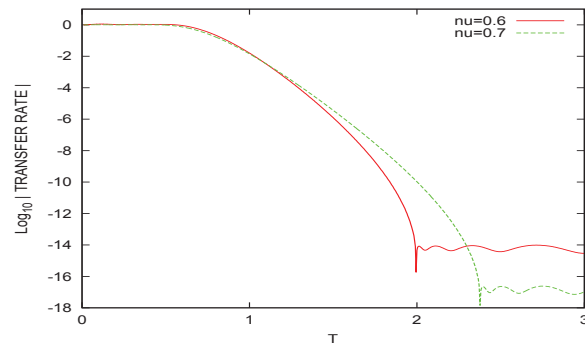


図 36 規格化 beta 関数の「拡張」($p=15, q=5$) (縦軸対数)

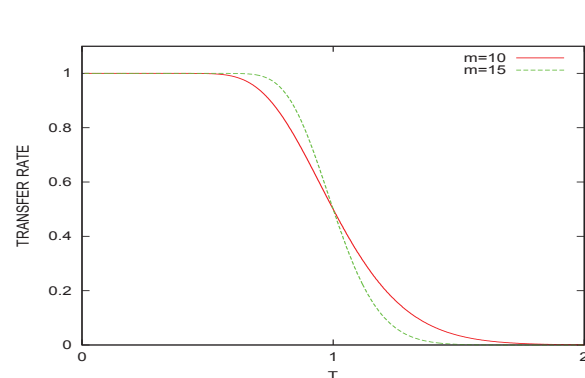


図 37 両端点近傍に選点条件を設定 (両軸真数)

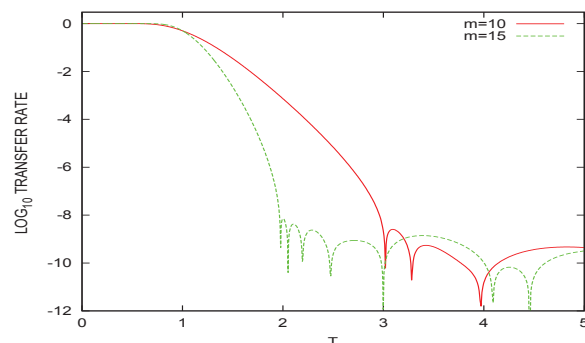


図 38 両端点近傍に選点条件を設定 (縦軸対数)

- ティングシステム (ACS35), Vol.4, No.4 (2011), pp.1-14.
- [3] NIST: Digital Library of Mathematical Functions, §8.17 Incomplete Beta Functions.
url="http://dlmf.nist.gov/8.17".