

Avanços em regularização de Tikhonov multiparamétrica para problemas discretos mal-postos.

Leonardo Silveira Borges (UNICAMP)
Maria Cristina C. Cunha (UNICAMP)
Fermín S. V. Bazán (UFSC)

UNICAMP
Departamento de Matemática Aplicada

03 de setembro, 2011

Estrutura

- 1 Introdução
- 2 Algoritmo de ponto fixo
- 3 GKB-FP: problemas de grande porte
- 4 Extensões de GKB-FP
- 5 FP para Tikhonov multi-paramétrico

Estrutura

- 1 **Introdução**
- 2 Algoritmo de ponto fixo
- 3 GKB-FP: problemas de grande porte
- 4 Extensões de GKB-FP
- 5 FP para Tikhonov multi-paramétrico

Problema discreto linear mal-posto

Problema discreto linear mal-posto:

$$f = \underset{f \in \mathbb{R}^n}{\operatorname{argmin}} \|g - Af\|_2^2 \quad (1)$$

$A \in \mathbb{R}^{m \times n}$, $m \geq n$, mal condicionada.

$g = g^{\text{exato}} + e \in \mathbb{R}^m$.

Consequência: Solução $f_{LS} = A^\dagger g$ sem utilidade.

Alternativa: Tikhonov propôs substituir (1) por

$$f_\lambda = \underset{f \in \mathbb{R}^n}{\operatorname{argmin}} \{ \|g - Af\|_2^2 + \lambda^2 \|Lf\|_2^2 \}$$

$\lambda > 0$: **Parâmetro de regularização.**

$L \in \mathbb{R}^{p \times n}$: **Matriz de regularização.**

Métodos para escolha de λ

Métodos clássicos:

- Princípio da Discrepância (Morozov, 1966)
- Validação Cruzada-Generalizada (Golub, 1979)
- Critério da Curva-L (Hansen, 1993)

Propostas mais recentes:

- Ponto-Fixo (Bazán, 2008, Bazán e Francisco, 2009)
- Validação Cruzada-Generalizada com peso (Chung, 2008)

Comparativo entre 20 métodos para escolha de λ : F. Bauer e M. A. Lukas, Comparing parameter choice methods for regularization of ill-posed problems, 2011.

Curva-L e o critério do produto mínimo

Norma do resíduo: $x(\lambda) = \|g - Af_\lambda\|_2^2$, função crescente

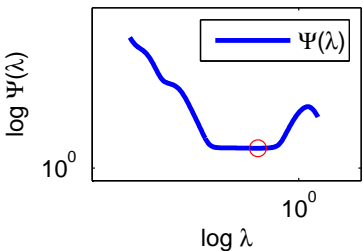
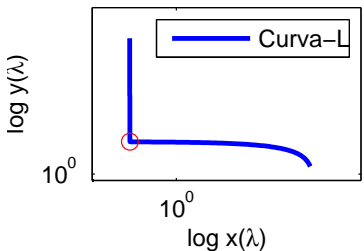
(semi-)Norma da solução: $y(\lambda) = \|Lf_\lambda\|_2^2$, função decrescente

Hansen, 1993: Encontrar o ponto de máxima curvatura de

$$\mathcal{L} = \{(a, b) \in \mathbb{R}^2 / a = \log x(\lambda), b = \log y(\lambda)\}$$

Regińska, 1996: Encontrar um minimizador local da função

$$\Psi(\lambda) = x(\lambda)y(\lambda)^\mu, \quad \mu > 0$$



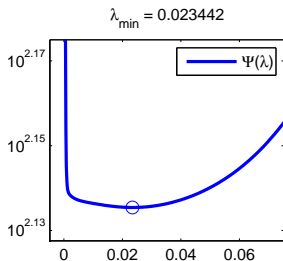
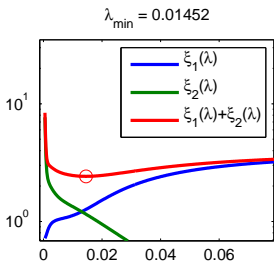
Justificativa do critério do produto mínimo

Solução regularizada: $f_\lambda = (A^T A + \lambda^2 L^T L)^{-1} A^T g = R_\lambda g$

Cota para o erro:

$$\begin{aligned} \|f^{\text{exato}} - f_\lambda\|_2 &= \|f^{\text{exato}} - R_\lambda g^{\text{exato}} - R_\lambda e\|_2 \\ &\leq \|f^{\text{exato}} - R_\lambda g^{\text{exato}}\|_2 + \|R_\lambda e\|_2 = \xi_1(\lambda) + \xi_2(\lambda) \end{aligned}$$

log do critério do produto: $\log \Psi(\lambda) = \log x(\lambda) + \mu \log y(\lambda)$



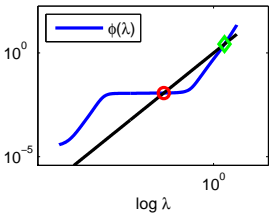
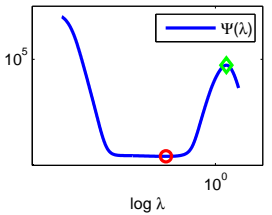
Estrutura

- 1 Introdução
- 2 Algoritmo de ponto fixo**
- 3 GKB-FP: problemas de grande porte
- 4 Extensões de GKB-FP
- 5 FP para Tikhonov multi-paramétrico

Ponto Fixo (FP) (Bazán, 2008 e Bazán e Francisco, 2009)

Pontos críticos de $\Psi(\lambda)$ estão associados aos pontos fixos de

$$\phi(\lambda; \mu) = \sqrt{\mu} \frac{\|g - Af_\lambda\|_2}{\|Lf_\lambda\|_2}$$



Algoritmo FP: $\lambda_0 > 0$: chute inicial
 $\lambda_{k+1} = \phi(\lambda_k; \mu), \quad k \geq 0$

▶ multi

Estrutura

- 1 Introdução
- 2 Algoritmo de ponto fixo
- 3 GKB-FP: problemas de grande porte**
- 4 Extensões de GKB-FP
- 5 FP para Tikhonov multi-paramétrico

Bidiagonalização de Golub-Kahan e LSQR

Bidiag. de Golub-Kahan: No passo k existem matrizes U_{k+1} e V_k com colunas ortonormais e $B_k \in \mathbb{R}^{(k+1) \times k}$ bidiag. inferior tal que

$$\begin{aligned}
 AV_k &= U_{k+1}B_k \\
 A^T U_{k+1} &= V_k B_k^T + \alpha_{k+1} v_{k+1} e_{k+1}^T
 \end{aligned}$$

Propriedade: $\text{span}(V_k) = \mathcal{K}_k(A^T A, A^T b)$

Algoritmo LSQR: Construir soluções $f_k = V_k y_k$ com

$$y_k = \underset{y \in \mathbb{R}^k}{\text{argmin}} \|\beta_1 e_1 - B_k y\|_2^2$$

em que $\beta_1 = \|g\|_2$.

LSQR e Tikhonov

No problema regularizado, construir soluções $f_{k,\lambda} = V_k y_{k,\lambda}$ com

$$y_{k,\lambda} = \operatorname{argmin}_{y \in \mathbb{R}^k} \{ \|\beta_1 e_1 - B_k y\|_2^2 + \lambda^2 \|y\|_2^2 \}$$

em que $\beta_1 = \|g\|_2$.

Propriedades: para $\lambda \geq 0$ fixo,

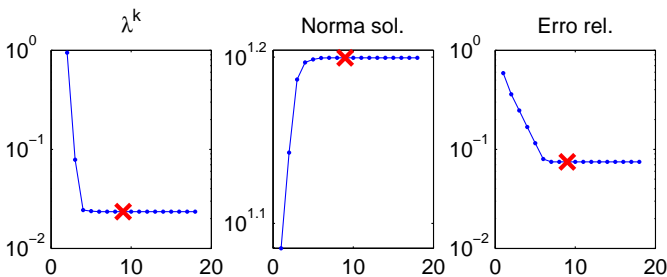
$\|\beta_1 e_1 - B_k y_{k,\lambda}\|_2$ é **decrecente** em k

$\|y_{k,\lambda}\|_2$ é **crescente** em k

GKB-FP, Estabilização da solução

Teorema 3

Se o maior ponto fixo convexo de $\phi(\lambda; 1)$ é encontrado no passo k , isto é, $\lambda^* = \phi^{(k)}(\lambda^*; 1)$ para algum k , então $f_{k, \lambda^*} = f_{k+1, \lambda^*} = \dots = f_{n, \lambda^*} \triangleq f_{\lambda^*}$ e como consequência, a norma do erro $\|f^{exato} - f_{j, \lambda^*}\|_2$ permanece constante para $k \leq j \leq n$.



GKB-FP para forma geral (Bazán e Borges, 2011)

Ideias gerais:

1. Transformar o problema na forma geral

$$f_\lambda = \operatorname{argmin}_{f \in \mathbb{R}^n} \{ \|g - Af\|_2^2 + \lambda^2 \|Lf\|_2^2 \}, \quad L \neq I$$

para a forma padrão

$$\bar{f}_\lambda = \operatorname{argmin}_{\bar{f} \in \mathbb{R}^n} \{ \|\bar{g} - \bar{A}\bar{f}\|_2^2 + \lambda^2 \|\bar{f}\|_2^2 \}$$

2. Aplicar GKB-FP no problema transformado na forma padrão.
3. Obtida solução \bar{f}_λ , calcular f_λ .

Obs: Transformação para forma padrão envolve cálculos com a matriz L^\dagger .

Regularizador L com estrutura especial

Considerar L com estrutura

$$L = \begin{bmatrix} I_{\tilde{n}} \otimes L_d \\ L_d \otimes I_{\tilde{n}} \end{bmatrix}$$

em que $L_d \in \mathbb{R}^{(\tilde{n}-d) \times \tilde{n}}$.

Propriedade 1: Estrutura especial de L permite usar a SVD de L_d .

$L_d = U_d \Sigma_d V_d^T$ e encontrar matriz L_D com tal que

$\|Lf\|_2 = \|L_D f\|_2$ com

$$L_D = D(V_d \otimes V_d)^T \quad \text{e} \quad D^2 = \Sigma_d^T \Sigma_d \otimes I + I \otimes \Sigma_d^T \Sigma_d$$

Propriedade 2: Redução do custo comp. para $L^\dagger f$ e $L^{\dagger T} f$

$$L_D^\dagger f = (V_d \otimes V_d) D^\dagger f \quad \text{e} \quad L_D^{\dagger T} f = D^\dagger (V_d \otimes V_d)^T f$$

PROJ-FP (Bazán e Borges, 2011)

Problema na **forma geral**

$$f_\lambda = \operatorname{argmin}_{f \in \mathbb{R}^n} \{ \|g - Af\|_2^2 + \lambda^2 \|Lf\|_2^2 \}$$

Ideia: Construir soluções $f_{k,\lambda} = V_k y_{k,\lambda}$ com

$$y_{k,\lambda} = \operatorname{argmin}_{y \in \mathbb{R}^k} \{ \|\beta_1 e_1 - B_k y\|_2^2 + \lambda^2 \|LV_k y\|_2^2 \}$$

Decomposição QR de LV_k , $LV_k = Q_k R_k$, implica

$$y_{k,\lambda} = \operatorname{argmin}_{y \in \mathbb{R}^k} \{ \|\beta_1 e_1 - B_k y\|_2^2 + \lambda^2 \|R_k y\|_2^2 \}$$

Propriedade: Fácil atualização da QR de LV_{k+1} .

G-LSQR (Bazán e Borges, 2011)

Ideia geral: Aplicar o algoritmo LSQR ao problema

$$\operatorname{argmin}_{\bar{f} \in \mathbb{R}^n} \|\bar{g} - \bar{A}\bar{f}\|_2^2$$

Critério de parada: Índice k^* tal que

$$k^* = \operatorname{argmin} \bar{\Psi}_k, \quad k \geq 0$$

em que

$$\bar{\Psi}_k = \|\bar{g} - \bar{A}\bar{f}_k\|_2 \|\bar{f}_k\|_2$$

Obs: Obtida solução \bar{f}_k , calcular f_k .

Exemplo - Restauração de imagem



Imagem: 512×512 pixels

Sistema linear: 262.144 variáveis

Nível ruído: 0,1%

Regularizador: $L = \begin{bmatrix} I \otimes \mathcal{L}_1 \\ \mathcal{L}_1 \otimes I \end{bmatrix}_{523264 \times 262144}$

Exemplo - Restauração de imagem

	L_D	PROJ-FP	G-LSQR
E	15,92%	12,53%	15,88%
tempo	18,8s	188,4s	17,6s
iterações	154	171	154
λ	0,1377	0,0114	—



Estrutura

- 1 Introdução
- 2 Algoritmo de ponto fixo
- 3 GKB-FP: problemas de grande porte
- 4 Extensões de GKB-FP
- 5 FP para Tikhonov multi-paramétrico**

Tikhonov Multi-paramétrico

O problema com vários parâmetros:

$$f_\lambda = \operatorname{argmin}_{f \in \mathbb{R}^n} \{ \|g - Af\|_2^2 + \lambda_1^2 \|L_1 f\|_2^2 + \cdots + \lambda_q^2 \|L_q f\|_2^2 \}$$

$A \in \mathbb{R}^{m \times n}$, $L_i \in \mathbb{R}^{p_i \times n}$ e $\lambda = (\lambda_1, \dots, \lambda_q) \in \mathbb{R}^q$, $\lambda_i > 0$.

Dificuldade: Falta de decomposição “GSVD” para (A, L_1, \dots, L_q) do tipo $A = U\Sigma X$, $L_i = V_i M_i X$, em que U , V_i são ortogonais, Σ e M_i diagonais e X não-singular.

Ideia: Tratar como

$$f_\lambda = \operatorname{argmin}_{f \in \mathbb{R}^n} \left\| \begin{pmatrix} g \\ 0 \\ \vdots \\ 0 \end{pmatrix} - \begin{pmatrix} A \\ \lambda_1 L_1 \\ \vdots \\ \lambda_q L_q \end{pmatrix} f \right\|_2^2$$

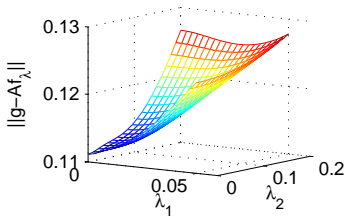
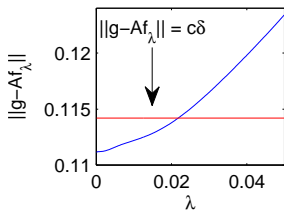
Discrepância para múltiplos parâmetros

Escolher $\lambda = (\lambda_1, \dots, \lambda_q)$ tal que a solução regularizada f_λ satisfaça o **Princípio da Discrepância**

$$\|g - Af_\lambda\|_2 = c\delta, \quad c \geq 1$$

em que $\|e\| = \|g - g^{\text{exato}}\|_2^2 \leq \delta$.

Interpretação geométrica usando 1 e 2 parâmetros:



Funções modelo (Lu e Pereverzev, 2011)

Considerar o caso

$$f_\lambda = \operatorname{argmin} \{ \|g - Af\|_2^2 + \lambda_1^2 \|L_1 f\|_2^2 + \lambda_2^2 \|f\|_2^2 \}$$

A equação da discrepância pode ser reescrita como

$$F(\lambda_1, \lambda_2) - \lambda_1^2 \partial_{\lambda_1} F(\lambda_1, \lambda_2) - \lambda_2^2 \partial_{\lambda_2} F(\lambda_1, \lambda_2) = c^2 \delta^2 \quad (2)$$

Ideia: aproximar $F(\lambda_1, \lambda_2)$ por uma função mais simples, $m(\lambda_1, \lambda_2)$, chamada de função modelo, que seja mais fácil de trabalhar.

$$m(\lambda_1, \lambda_2) = \|g\|_2^2 + \frac{C}{\lambda_1^2} + \frac{D}{T + \lambda_2^2}$$

Combinação linear (Brezinski, et al., 2003)

Trabalhar com q subproblemas uniparamétricos

$$f_{\lambda_i} = \operatorname{argmin}_{f \in \mathbb{R}^n} \{ \|g - Af\|_2^2 + q\lambda_i^2 \|L_i f\|_2^2 \}$$

Obter solução f_λ como **combinação linear** das soluções f_{λ_i}

$$f_\lambda = f_\lambda(\alpha) = \sum_{i=1}^q \alpha_i f_{\lambda_i}, \quad \text{s. a } \sum_{i=1}^q \alpha_i = 1$$

Critério do produto mínimo - Ponto Fixo

Sejam

$$x(\lambda) = \|g - Af_\lambda\|_2^2, \quad y_i(\lambda) = \|L_i f_\lambda\|_2^2$$

Escolher como parâmetro de regularização o ponto $\lambda \in \mathbb{R}^q$ minimizador local de

$$\Psi(\lambda) = x(\lambda)y_1(\lambda)^{\mu_1} \cdots y_q(\lambda)^{\mu_q}, \quad \mu_i > 0$$

Pontos extremos de $\Psi(\lambda)$ estão associados aos pontos fixos de

$$\Phi(\lambda) = (\phi_1(\lambda; \mu_1), \dots, \phi_q(\lambda; \mu_q))$$

em que

$$\phi_i(\lambda; \mu_i) = \sqrt{\mu_i} \frac{\|g - Af_\lambda\|_2}{\|L_i f_\lambda\|_2}$$

Pontos extremos de $\Psi(\lambda)$

Do cálculo, λ é ponto extremo se $\nabla\Psi(\lambda) = 0$, então

$$\nabla\Psi(\lambda) = y_1^{\mu_1}(\lambda) \cdots y_q^{\mu_q}(\lambda) J \begin{pmatrix} \mu_1 \frac{x(\lambda)}{y_1(\lambda)} - \lambda_1^2 \\ \vdots \\ \mu_q \frac{x(\lambda)}{y_q(\lambda)} - \lambda_q^2 \end{pmatrix} = 0$$

em que

$$J = \begin{bmatrix} \frac{\partial y_1}{\partial \lambda_1} & \cdots & \frac{\partial y_q}{\partial \lambda_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial \lambda_q} & \cdots & \frac{\partial y_q}{\partial \lambda_q} \end{bmatrix}, \quad \text{n\~{a}o-singular}$$

Algoritmo

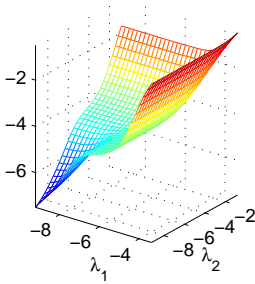
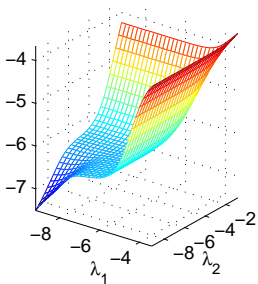
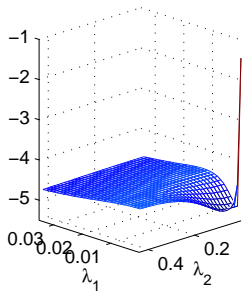
1. $\lambda^{(0)} = (\lambda_1^{(0)}, \dots, \lambda_q^{(0)}) > 0$
2. $\lambda^{(k+1)} = (\lambda_1^{(k+1)}, \dots, \lambda_q^{(k+1)}) = \Phi(\lambda^{(k)})$, isto é,
$$\lambda_i^{(k+1)} = \phi_i(\lambda^{(k)}; \mu_i), \quad i = 1, \dots, q$$

em que

$$\phi_i(\lambda; \mu_i) = \sqrt{\mu_i} \frac{\|g - Af_\lambda\|_2}{\|L_i f_\lambda\|_2}$$

Superfícies $\Psi(\lambda)$, $\phi_1(\lambda)$ e $\phi_2(\lambda)$

Exemplo de superfícies $\Psi(\lambda)$, $\phi_1(\lambda)$ e $\phi_2(\lambda)$



▶ FP

Exemplo - Equação integral

Discretização da transformada inversa de Laplace (equação integral de Fredholm de primeira espécie).

$$\int_0^{\infty} K(x, y)f(y)dy = g(x), \quad 0 \leq x < \infty$$

em que o **Núcleo** $K(x, y)$, **Solução** $f(y)$ e **Lado direito** $g(x)$ são dados por

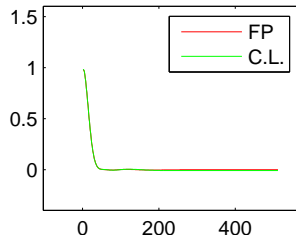
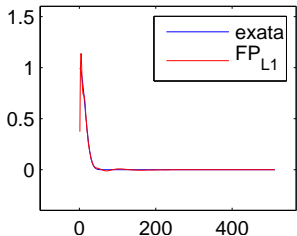
$$\begin{aligned}
 K(x, y) &= \exp(-xy) \\
 f(y) &= \exp(-y/2) \\
 g(x) &= \frac{1}{x + 1/2}
 \end{aligned}$$

Exemplo - Equação integral

Sistema linear: 512 variáveis








Nível ruído: 1% Regularizadores: $L_1 = I$ e $L_2 = \mathcal{L}_1$.

	FP $_{L_1}$	FP $_{L_2}$	FP	Comb. Linear	Discrep.
E	20,42%	4,51%	1,19%	4,50%	3,17%
iterações	4	4	6	–	7
λ_1	0,0281	0,5571	0,0276	0,0281	0,0389
λ_2	–	–	0,5581	0,5571	0,0008



Trabalhos futuros

- Análise de erro
- Possibilidade de estender FP para o caso não-linear

-  F. S. V. Bazán, *Fixed-point iterations in determining the Tikhonov regularization parameter*, Inv. Problems, 24, 2008.
-  F. S. V. Bazán e J. B. FRANCISCO, *Improved Fixed-point algorithm for determining the Tikhonov regularization parameter*, Inv. Problems, 2009.
-  F. S. V. Bazán e L. S. Borges, *GKB-FP: an algorithm for large-scale discrete ill-posed problems*, BIT, 50, p. 481-507, 2010.
-  F. S. V. Bazán e L. S. Borges, *Extension of GKB-FP algorithm to general-form Tikhonov regularization*, submetido.
-  C. Brezinski, M. Redivo-Zaglia, G. Rodriguez e S. Seatzu, *Multi-parameter regularization techniques for ill-conditioned linear systems*, Numerische Mathematik, 94, p. 203-228, 2003.
-  S. Lu e S. V. Pereverzev, *Multi-parameter regularization and its numerical realization*, Numerische Mathematik, 118, p. 1-31, 2011.
-  F. Bauer e M. A. Lukas, *Comparing parameter choice methods for regularization of ill-posed problems*, Math. and Comp. in Simulation, 81, p. 1795-1841, 2011.



V. A. Morozov, *On the solution of functional equations by the method of regularization*, Soviet Math. Dokl., 7, pp. 414-417, 1966.



G. H. Golub, M. T. Heath e G. Wahba, *Generalized cross-validation as a method for choosing a good ridge parameter*, Technometrics, 21, pp. 215-223, 1979.



C. C. Paige e M. A. Saunders, *LSQR: An algorithm for sparse linear equations and sparse least squares*, ACM Trans. Math. Softw, Vol. 8, No. 1, pp 43-71, 1982.



P. C. Hansen, *Rank-deficient and discrete ill-posed problems*, SIAM Philadelphia, PA, 1998.



J. Chung, J. G. Nagy, e D. P. O'Leary, *A Weighted-GCV Method for Lanczos-Hybrid Regularization*, Electronic Transaction on Numerical Analysis, 28, pp 149-167, 2008.