A Positive BB-Like Stepsize and An Extension for Symmetric Linear Systems

Yu-Hong Dai

Academy of Mathematics and Systems Science,
Chinese Academy of Sciences

Joint with M. Al-Baali and Xiaoqi Yang

Peking University, 20140903

Outline

Section I. Introduction

Unconstrained Optimization

$$\min \quad f(\mathbf{x}), \quad \mathbf{x} \in \mathcal{R}^n$$

Convex Quadratic Minimization

$$\min \quad Q(\mathbf{x}) := \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{b}^T\mathbf{x}, \quad \mathbf{x} \in \mathcal{R}^n$$

Linear System

$$A\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in \mathcal{R}^n$$

Steepest Descent Method (Cauchy 1847)

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \, \mathbf{g}_k$$

$$\alpha_k = \arg\min_{\alpha \geq 0} f(\mathbf{x}_k - \alpha \, \mathbf{g}_k)$$

- Fast during early several iterations
- Linear Convergence

$$\|\mathbf{g}_k\|_2 \approx \left(\frac{\kappa - 1}{\kappa + 1} \right)^k, \quad \kappa = \text{cond}\left(\nabla^2 f(\mathbf{x}^*)\right)$$

- Zigzagging

Barzilai-Borwein (1988)

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{x}_k - \alpha_k \, \mathbf{g}_k \\ &= \mathbf{x}_k - D_k^{-1} \, \mathbf{g}_k \end{aligned}$$

$$D_k = \arg \min_{D=\alpha^{-1} I} \| D_k \mathbf{s}_{k-1} - \mathbf{y}_{k-1} \|_2^2$$

$$(\mathbf{s}_{k-1} = \mathbf{x}_k - \mathbf{x}_{k-1}, \quad \mathbf{y}_{k-1} = \mathbf{g}_k - \mathbf{g}_{k-1})$$

$$\Rightarrow \quad \alpha_k^{BB1} = \frac{\mathbf{s}_{k-1}^T \mathbf{s}_{k-1}}{\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}}$$

Similarly,

$$\alpha_k^{BB2} = \frac{\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}}{\mathbf{y}_{k-1}^T \mathbf{y}_{k-1}}$$

Fletcher (2005), “On the Barzilai-Borwein method":

$$\triangle \mathbf{u} = -f, \quad \mathbf{u} \in [0,1]^3$$

$$f = x(x-1)y(y-1)z(z-1)w(x,y,z)$$

$$w = \exp\Big(-\frac{1}{2}\sigma^2((x-\alpha)^2+(y-\beta)^2+(z-\gamma)^2)\Big)$$

$$A\mathbf{u} = \mathbf{b}, \quad n = 10^6$$

$$\Big(\Leftrightarrow \min \frac{1}{2}\mathbf{u}^T A\mathbf{u} - \mathbf{b}^T\mathbf{u}\Big)$$

$$\mathbf{u}_1 = 0, \quad \|\mathbf{g}_k\|_2 \le 10^{-6}\|\mathbf{g}_1\|_2$$

Numerical Results

$(\sigma, \alpha, \beta, \gamma)$		BB	CG
(20, 0.5, 0.5, 0.5)	double	543(859)	162(178)
	single	462(964)	254(387)
(50, 0.4, 0.7, 0.5)	double	640(1009)	285(306)
	single	310(645)	290(443)

But SD: 2000, $\frac{\|\mathbf{g}_{2000}\|}{\|\mathbf{g}_1\|} = 0.18$!

Scholar google **BB**:

806 times (by Jan 5, 2014)

Scholar google **GPSR** by Figueiredo, Wright and Nowak (2007):

1310 times (by Jan 5, 2014)

Efficiency Evidences of BB for Quadratic Minimization

- Barzilai-Borwein (1988)

  $n = 2$, R-superlinear

  $$\left(\alpha_{k_{i_1}}^{-1} \to \lambda_1, \quad \alpha_{k_{i_2}}^{-1} \to \lambda_2\right)$$

- Dai & Fletcher (2005)

  $n = 3$, R-superlinear

- Dai & Fletcher (2005)

  Cyclic SD method, $m \geq \frac{n}{2} + 1$, R-superlinear

In theory, how to show that BB is better than SD for any-dimensional quadratic functions?

Quadratic Termination of Gradient Method

$$\begin{aligned} \mathbf{g}_{k+1} &= \mathbf{g}_k - \alpha_k A \mathbf{g}_k \\ &= (I - \alpha_k A)\mathbf{g}_k \\ &= \left[\prod_{j=1}^{k} (1 - \alpha_j A) \right] \mathbf{g}_1 \end{aligned}$$

Assuming that

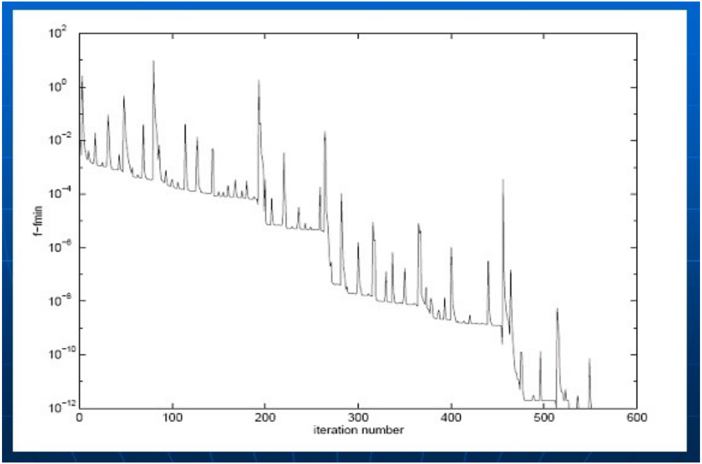
$$\lambda(A) = \{\lambda_1, \lambda_2, ..., \lambda_n\}$$

by the Caylay-Hamilton theorem, we must have $\mathbf{g}_{n+1} = 0$ if

$$\left\{\alpha_k : \; k = 1, ..., n\right\} = \left\{\lambda_k^{-1} : \; k = 1, ..., n\right\}$$

This property was first due to Yan-Lian Lai (1983).

A Typical Nonmonotone Performance of BB



For any dimensional strictly convex quadratics

- Raydan (1993): global convergence

- Dai & Liao (2002): *R*-linear convergence

  We can then show that the BB stepsize can be asymptotically accepted by the nonmonotone line search in the context of unconstrained optimization. This is a property similar to quasi-Newton methods where the stepsize $\alpha_k = 1$ is usually firstly tried by the Wolfe line search and it will gradually accepted.

Gobalization Technique for General Functions

Raydan (1997): GLL nonmonotone line search

$$f(x_k - \alpha \mathbf{g}_k) \le f_{ref} - \delta \alpha \|\mathbf{g}_k\|^2, \quad f_{ref} = \max_{j=1,\dots,m} f_{k-j}$$

Dai & Zhang (2001): Adaptive nonmonotone line search

$$\text{Initialization}: \ f_{ref} = +\infty, \quad H \in [4, 10]$$

$$\begin{array}{l} \text{If } f_k \le f_{best} \\ \quad f_{best} = f_k, \ \ f_c = f_k, \ \ h = 0; \\ \text{Else} \\ \quad f_c = \max\{f_c, f_k\}, \ \ h = h + 1 \\ \quad \text{if } h = H, \ \ f_{ref} = f_c, \ \ search, \ \ f_c = f_k, \ \ h = 0 \end{array}$$

Section II. A Positive BB-Like Stepsize

Motivation

- What to do if the BB stepsize

$$\alpha_k^{BB1} = \frac{\mathbf{s}_{k-1}^T \mathbf{s}_{k-1}}{\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}} \quad \text{or} \quad \alpha_k^{BB2} = \frac{\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}}{\mathbf{y}_{k-1}^T \mathbf{y}_{k-1}}$$

  is very small or even negative? Project it onto the interval

$$\left[\alpha_k^{\min},\ \alpha_k^{\max}\right]?$$

  How to choose $\alpha_k^{\min}$ (and $\alpha_k^{\max}$)? 10^{-30}, 10^{-8}, 10^{-5}, ...

- For a symmetric but not necessarily positive definite linear system

$$A\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in \mathcal{R}^n,$$

  how to approximate the (inverse) Jacobian matrix by the form αI, in which case it may have negative eigenvalues?

The New positive stepsize

- The New positive stepsize

$$\alpha_k = \frac{\|\mathbf{s}_{k-1}\|}{\|\mathbf{y}_{k-1}\|} \tag{1}$$

Mentioned in several previous occasions, but not been carefully studied [eg., Dai & Yuan (2001), Dai (2003), Dai & Yang (2006), Mehiddin Al-Baali (2007)]

- Property 1: Geometry mean

$$\alpha_k = \sqrt{\alpha_k^{BB1} \cdot \alpha_k^{BB2}} \tag{2}$$

The New positive stepsize (Cond.)

Propery 2: Certain quasi-Newton property

- Two features of $\nabla^2 f(\mathbf{x}_k)$

$$\mathbf{s}_{k-1}^T \nabla^2 f(\mathbf{x}_k) \mathbf{s}_{k-1} \approx \mathbf{s}_{k-1}^T \mathbf{y}_{k-1} \quad (3)$$

$$\mathbf{y}_{k-1}^T \nabla^2 f(\mathbf{x}_k)^{-1} \mathbf{y}_{k-1} \approx \mathbf{s}_{k-1}^T \mathbf{y}_{k-1} \quad (4)$$

- Approximation

$\nabla^2 f(x_k)^{-1} \leftarrow H = \alpha I, \ \nabla^2 f(x_k) \leftarrow H^{-1} = \alpha^{-1} I$

$$\alpha_k = \arg \min_{H=\alpha I \succeq 0} \left| \mathbf{s}_{k-1}^T H^{-1} \mathbf{s}_{k-1} + \mathbf{y}_{k-1}^T H \mathbf{y}_{k-1} - 2\mathbf{s}_{k-1}^T \mathbf{y}_{k-1} \right|,$$

- Property 3: One-retard extension of [Dai & Yang, 2006]

$$\alpha_k^{DY} = \frac{\|\mathbf{g}_k\|}{\|A\mathbf{g}_k\|} \tag{5}$$

The stepsize (5) is shown to tend to some optimal stepsize:

$$\liminf_{k\to\infty} \alpha_k^{DY} = \frac{2}{\lambda_1 + \lambda_n} := \arg\min_{\alpha\geq 0} \|I - \alpha A\|. \tag{6}$$

Both the solution and the minimal/maximal eigenpairs can simultaneously obtained (One stone Two birds).

Section III. Analysis of The New Method

Some notations

Assume that

$$A = \begin{bmatrix} 1 & \\ & \lambda \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \ \lambda > 1$$

- Denote $\mathbf{g}_k = (g_k^{(1)}, g_k^{(2)})^{\mathsf{T}}$
- Assumption 1

$$\lambda > 1 \tag{7}$$

- Assumption 2

$$g_1^{(i)} \neq 0, \ \ g_2^{(i)} \neq 0, \ \ i = 1, 2 \tag{8}$$

- Define

$$q_k = \frac{\left(g_k^{(1)}\right)^2}{\left(g_k^{(2)}\right)^2} \tag{9}$$

Some basic relations

$$\alpha_k = \frac{\|\mathbf{s}_{k-1}\|}{\|\mathbf{y}_{k-1}\|} = \frac{\|\mathbf{g}_{k-1}\|}{\|A\mathbf{g}_{k-1}\|} = \frac{\sqrt{1+q_{k-1}}}{\sqrt{\lambda^2+q_{k-1}}} \tag{10}$$

$$\mathbf{g}_{k+1} = (I - \alpha_k A)\mathbf{g}_k \tag{11}$$

$$\begin{cases} g_{k+1}^{(1)} = (1-\alpha_k)g_k^{(1)} \\ g_{k+1}^{(2)} = (1-\lambda\alpha_k)g_k^{(2)} \end{cases} \Longrightarrow \begin{cases} g_{k+1}^{(1)} = \frac{\sqrt{\lambda^2+q_{k-1}}-\sqrt{1+q_{k-1}}}{\sqrt{\lambda^2+q_{k-1}}} g_k^{(1)} \\ g_{k+1}^{(2)} = \frac{\sqrt{\lambda^2+q_{k-1}}-\lambda\sqrt{1+q_{k-1}}}{\sqrt{\lambda^2+q_{k-1}}} g_k^{(2)} \end{cases} \tag{12}$$

Recurrence relation of q_k

$$
\begin{aligned}
q_{k+1} &= \left(\frac{\sqrt{\lambda^2 + q_{k-1}} - \sqrt{1 + q_{k-1}}}{\sqrt{\lambda^2 + q_{k-1}} - \lambda\sqrt{1 + q_{k-1}}} \right)^2 q_k \\
&= \left(\frac{(\sqrt{\lambda^2 + q_{k-1}} - \sqrt{1 + q_{k-1}})(\sqrt{\lambda^2 + q_{k-1}} + \lambda\sqrt{1 + q_{k-1}})}{(\lambda^2 - 1)q_{k-1}} \right)^2 q_k \\
&= \left(\frac{\lambda - q_{k-1} + \sqrt{\tau(q_{k-1})}}{\lambda - 1} \right)^2 \frac{q_k}{q_{k-1}^2}, \qquad (13)
\end{aligned}
$$

where

$$\tau(w) = (1 + w)(\lambda^2 + w), \ \ w \geq 0 \qquad (14)$$

$$h(w) = \frac{\lambda - w + \sqrt{\tau(w)}}{\lambda + 1}, \ \ w \geq 0 \qquad (15)$$

Define $M_k = \log q_k$. Then we obtain

$$M_{k+1} = M_k - 2M_{k-1} + 2\log\left(h(q_{k-1})\right) \qquad (16)$$

The difficulty:

Previously, for the BB1 or BB2 method, we can get the linear recurrence relation

$$M_{k+1} = M_k - 2M_{k-1}.$$

But now we have got a nonlinear recurrence relation.

Superlinear convergence

- Lower and upper bounds of $h(w)$

$$h(w) \in \left[\frac{2\lambda}{\lambda+1}, \frac{\lambda+1}{2}\right), \quad w \geq 0 \tag{17}$$

- Divergence of a subsequence of $\{M_k\}$
  - $\xi_k = M_k + (\gamma - 1)M_{k-1}$, $\gamma^2 - \gamma + 2 = 0$
  - Denote $c_2 = 2\log\frac{\lambda+1}{2}$ and assume that

$$c_1 := |\xi_2| - c_2 > 0 \tag{18}$$

  - Relation

$$|\xi_k| \geq c_1 2^{k-2} + c_2, \text{ for all } k \geq 2 \tag{19}$$

- Divergence

$$|\xi_k| \leq |M_k| + 2|M_{k-1}| \leq 3\max\{|M_k|, |M_{k-1}|\}$$

$$\Longrightarrow \max\{|M_k|, |M_{k-1}|\} \geq \frac{1}{3}(c_1 2^{k-2} + c_2) \tag{20}$$

Superlinear convergence (Cond.)

- Two subsequences of $\{M_k\}$ which tend to $+\infty$ and $-\infty$

$$\max_{-1\le i\le 3} M_{k+i} \ge \frac{1}{3}c_1 2^{k-2} - 2c_2 \tag{21}$$

$$\min_{-1\le i\le 3} M_{k+i} \le -\frac{1}{3}c_1 2^{k-2} + 2c_2 \tag{22}$$

Proof.

- Recursive relations

$$M_{k+1} = M_k - 2M_{k-1} + 2\log h(q_{k-1}) \tag{23}$$

$$M_{k+2} = -M_k - 2M_{k-1} + 2\log h(q_k) + 2\log h(q_{k-1}) \tag{24}$$

- $M_{k-i} \ge \frac{1}{3}(c_1 2^{k-2} + c_2)$ holds for some $i = 0$ or 1
- $M_{k-i} \le -\frac{1}{3}(c_1 2^{k-2} + c_2)$ holds for some $i = 0$ or 1
  - If $M_{k-i+1} \ge 0$, then $M_{k-i+2} \ge \frac{2}{3}(c_1 2^{k-2} + c_2) - 2c_2$
  - If $M_{k-i+1} \le 0$, then $M_{k-i+3} \ge \frac{2}{3}(c_1 2^{k-2} + c_2) - 2c_2$.

Superlinear convergence

Under assumptions (7), (8) and (18), $\{\|\mathbf{g}_k\|\} \to 0$, R-superlinear

Proof.

- Basic
$$|g_{k+1}^{(i)}| \le (\lambda - 1)|g_k^{(i)}|, \quad i = 1, 2 \tag{25}$$
- $|g_{k+1}^{(2)}| \le (\lambda - 1)^5 \exp\left(-\frac{1}{3}c_1 2^{k-2} + 2c_2\right) |g_k^{(2)}|$
  - $|g_{k+1}^{(2)}| < (\lambda - 1) q_{k-1} |g_k^{(2)}|$
  - $|g_{k+5}^{(2)}| \le (\lambda - 1)^5 \left(\min_{-1 \le i \le 3} q_{k+i}\right) |g_k^{(2)}|$
- $|g_{k+5}^{(1)}| \le \frac{1}{2}(\lambda + 1)(\lambda - 1)^5 \exp\left(-\frac{1}{3}c_1 2^{k-2} + 2c_2\right) |g_k^{(2)}|$
  - $|g_{k+1}^{(1)}| < \frac{\lambda^2 - 1}{q_{k-1}} |g_k^{(1)}|$
- $\|\mathbf{g}_{k+5}\| \le \frac{1}{2}(\lambda + 1)(\lambda - 1)^5 \exp\left(-\frac{1}{3}c_1 2^{k-2} + 2c_2\right) \|\mathbf{g}_k\|$

R-linear convergence

- $n \geq 3$, R-linear convergence, Dai (2003)

$$\alpha_k = \frac{\|\mathbf{s}_{\nu(k)+1}\|}{\|\mathbf{y}_{\nu(k)+1}\|}, \ \nu(k) \in \{k, k-1, \max\{k-m+1, 1\}\} \tag{26}$$

- General nonlinear function

$$\begin{aligned} \bar{\alpha}_k^{BB1} &= \max\left\{\alpha_k^{BB1}, \alpha_k^{new}\right\} \\ &= \max\left\{\frac{\mathbf{s}_{k-1}^T \mathbf{s}_{k-1}}{\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}}, \frac{\|\mathbf{s}_{k-1}\|}{\|\mathbf{y}_{k-1}\|}\right\} \end{aligned} \tag{27}$$

Section IV. An Extension for Symmetric Linear Systems

Symmetric Linear Systems

Consider the symmetric linear system

$$A\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in R^n$$

where $A = A^T$ nonsingular.

- Stepsize

$$\alpha_k = \mathrm{sign}(\mathbf{s}_{k-1}^T \mathbf{y}_{k-1}) \frac{\|\mathbf{s}_{k-1}\|}{\|\mathbf{y}_{k-1}\|} \tag{28}$$

  where

$$\mathrm{sign}(a) = \begin{cases} 1, & a \geq 0 \\ -1, & a < 0 \end{cases}$$

- Test instances

Ex. 1 $v = (-1)^i i,\ i = 1 : n$

Ex. 2 $v = -n/2 + n * rand(n, 1)$

Ex. 3 $v = randn(n, 1)$

Ex. 4 $v1 = -1 + (-a + 1) * rand(n_1, 1)$
$v2 = -a + 2a * rand(n_1, 1)$
$v3 = a + (1 - a) * rand(n - 2n_1, 1)$
$v = (v1; v2; v3)$ $n_1 = floor(n/3)$ and $a \in (0, 1)$

n	10	20	30	40	50
Ex. 1, tol = 1e-6					
BB1	672, 1e-06	2804, 1e-06	6415, 1e-06	11501, 1e-06	18061, 1e-06
BB2	210, 7e-07	638, 1e-06	1059, 9e-07	1880, 9e-07	2936, 1e-06
(28)	146, 7e-07	413, 8e-07	583, 1e-06	821, 7e-07	790, 9e-07
Ex. 2, tol = 1e-6					
BB1	267, 7e-07	1797, 1e-06	20000, 5e+06	20000, 2e+01	20000, 7e-02
BB2	129, 8e-07	383, 9e-07	7750, 9e-07	5907, 1e-06	7272, 9e-07
(28)	118, 1e-06	193, 9e-07	2207, 9e-07	1977, 9e-07	2412, 1e-06
Ex. 3, tol = 1e-6					
BB1	5750, 1e-06	20000, 3e+68	14037, 1e-06	20000, 1e-01	20000, 8e+02
BB2	371, 8e-07	20000, 1e-05	698, 1e-06	8019, 9e-07	16877, 1e-06
(28)	294, 4e-07	5562, 1e-06	420, 5e-07	2969, 9e-07	3517, 7e-07
Ex. 4, tol = 1e-3, $a = 0.001$					
BB1	111, 1e-03	20000, 2e-03	20000, 1e-01	20000, 3e-03	20000, 6e+98
BB2	55, 9e-04	4465, 1e-03	20000, 1e-03	20000, 1e-03	20000, 2e-03
(28)	60, 9e-04	656, 9e-04	5325, 1e-03	1074, 1e-03	5702, 1e-03
Ex. 4, tol = 1e-3, $a = 0.01$					
BB1	20000, 4e-02	20000, 2e+12	20000, 1e+75	20000, 5e+03	20000, 4e+02
BB2	1845, 1e-03	6425, 1e-03	18801, 1e-03	18928, 1e-03	20000, 1e-03
(28)	1002, 1e-03	851, 9e-04	2378, 1e-03	2448, 1e-03	2679, 1e-03

$\text{sign}(\mathbf{s}_{k-1}^T \mathbf{y}_{k-1})$ is necessary

- Let us choose

$$A = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- Constant stepsize

$$\alpha_k = \frac{\|\mathbf{s}_{k-1}\|}{\|y_{k-1}\|} = 1$$

$$\begin{pmatrix} g_{k+1}^{(1)} \\ g_{k+1}^{(2)} \end{pmatrix} = \begin{pmatrix} (1-\alpha_k) g_k^{(1)} \\ (1+\alpha_k) g_k^{(2)} \end{pmatrix} = \begin{pmatrix} 0 \\ 2 g_k^{(2)} \end{pmatrix}$$

- $\|\mathbf{g}_k\|$ goes to infinity at a fast rate.

Superlinear convergence, $n = 2$

- Stepsize

$$\alpha_k = \text{sign}\left(\left(g_{k-1}^{(1)}\right)^2 - \left(g_{k-1}^{(2)}\right)^2\right) \frac{\sqrt{1+q_{k-1}}}{\sqrt{\lambda^2+q_{k-1}}} = \text{sign}\,(q_{k-1}-\lambda)\,\frac{\sqrt{1+q_{k-1}}}{\sqrt{\lambda^2+q_{k-1}}} \tag{29}$$

- If $q_{k-1} \geq \lambda$, there holds

$$\begin{cases} g_{k+1}^{(1)} = \frac{\sqrt{\lambda^2+q_{k-1}} - \sqrt{1+q_{k-1}}}{\sqrt{\lambda^2+q_{k-1}}} g_k^{(1)} \\ g_{k+1}^{(2)} = \frac{\sqrt{\lambda^2+q_{k-1}} + \lambda\sqrt{1+q_{k-1}}}{\sqrt{\lambda^2+q_{k-1}}} g_k^{(2)} \end{cases} \tag{30}$$

- If $q_{k-1} < \lambda$, there holds

$$\begin{cases} g_{k+1}^{(1)} = \frac{\sqrt{\lambda^2+q_{k-1}} + \sqrt{1+q_{k-1}}}{\sqrt{\lambda^2+q_{k-1}}} g_k^{(1)} \\ g_{k+1}^{(2)} = \frac{\sqrt{\lambda^2+q_{k-1}} - \lambda\sqrt{1+q_{k-1}}}{\sqrt{\lambda^2+q_{k-1}}} g_k^{(2)} \end{cases} \tag{31}$$

Superlinear convergence, $n = 2$ (Cond.)

- Recursive relation of q_k

$$q_{k+1} = h(q_{k-1})^2 \frac{q_k}{q_{k-1}^2} \tag{32}$$

$$h(w) = \begin{cases} \frac{\sqrt{\tau(w)}-(\lambda+w)}{\lambda-1}, & w \in [\lambda, +\infty) \\ \frac{\sqrt{\tau(w)}+(\lambda+w)}{\lambda-1}, & w \in [0, \lambda) \end{cases} \tag{33}$$

- $h(w)$ has lower and upper bounds

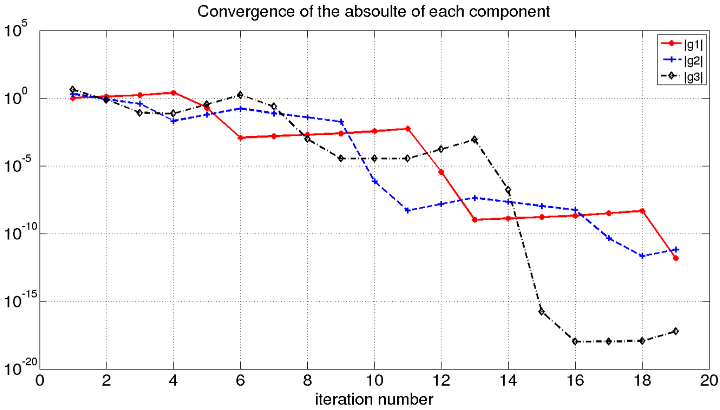
$$h(w) \in \begin{cases} \left[\frac{\sqrt{\lambda}-1}{\sqrt{\lambda}+1}\sqrt{\lambda}, \frac{\lambda-1}{2}\right), & w \in [\lambda, +\infty) \\ \left[\frac{2\lambda}{\lambda-1}, \frac{\sqrt{\lambda}+1}{\sqrt{\lambda}-1}\sqrt{\lambda}\right), & w \in [0, \lambda) \end{cases} \tag{34}$$

- Recursive relation of M_k

$$M_{k+1} = M_k - 2M_{k-1} + 2\log\left(h(q_{k-1})\right) \tag{35}$$

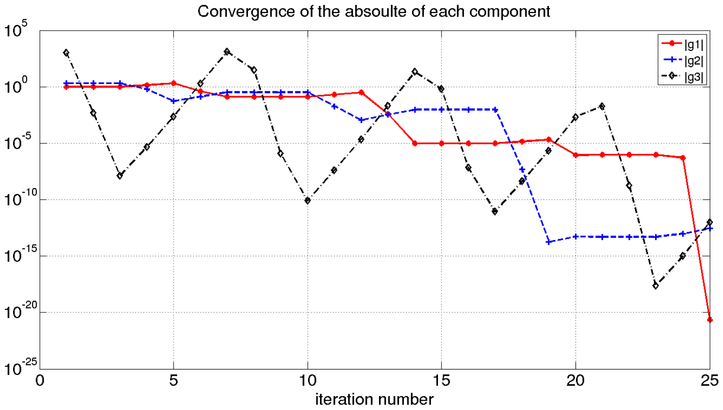
$n \geq 3$?

$A = \text{diag}(-1; 2; 4)$, $b = \text{zeros}(3, 1)$, $x = \text{ones}(3, 1)$



$n \geq 3$?

$A = \text{diag}(-1; 2; 1000)$, $b = \text{zeros}(3, 1)$, $x = \text{ones}(3, 1)$



Section V. Some Discussions

Some Discussions

- Optimization problem
  - How to relax $\xi_2 > c_2$?
  - How to Show the efficiency?
- Linear system of equations
  - *R*-linear convergence?
  - More efficient stepsize?
  - Non-symmetric problem?
- Nonlinear system of equations
  - How to improve the method by Cruz et al. (2006)?

Thank You!