الگوریتم های چولسکی تغییر یافته

رضا پیشکو

دانشگاه صنعتی شریف

ارائه سمینار محاسبات علمی
۱۴۰۲/۳/۸

مقدمه

- A متقارن و احتمالا غیر معین
- هدف : پیدا کردن $\hat{A} = A + E$ که مثبت معین باشد برای محاسبه E باید ۴ خاصیت زیر برقرار باشند.

1. اگر A به اندازه کافی مثبت معین باشد، آنگاه $E = ۰$.
2. اگر A مثبت معین نباشد، آنگاه $||E||$ خیلی بزرگ تر از $\inf\{||\Delta A|| : A + \Delta A$ نباشد.
3. ماتریس $A + E$ خوش تعریف باشد.
4. هزینه الگوریتم فقط به اندازه یک ضریب کوچکی از $n^۲$ بزرگ تر از فاکتورسازی استاندارد چولسکی $(\frac{۱}{۳}n^۳ + O(n^۲))$ باشد.

مقدمه

$$A = \mathrm{diag}\left(d_۱, d_۲, \ldots, d_n\right)$$

تخمین مثبت معین ماتریس A را تعریف کنیم :

$$A + E = \mathrm{diag}\left(\hat{d}_۱, \hat{d}_۲, \ldots, \hat{d}_n\right)$$

۱. الگوریتم نوع ۱ :

$$\hat{d}_k = \max\{|d_k|, \delta\}$$

۲. الگوریتم نوع ۲

$$\hat{d}_k = \max\{d_k, \delta\}$$

تجزیه های ماتریسی

- $PAP^T = LDL^T$ به طوری که D قطری باشد
- $PAP^T = LBL^T$ به طوری که B قطری_بلوکی با بلوک های ۱ یا ۲ بعدی باشد
- $PAP^T = LTL^T$ به طوری که T ماتریس سه قطری باشد.

پیچیدگی محاسباتی

الگوریتم	خاصیت ۴	الگوریتم	خاصیت ۴
LDL^T		LBL^T	
GMW۸۱	$O(n^{۲})$	MS۷۹	$\leq O(n^{۳})$
SE۹۰	$O(n^{۲})$	CH۹۸	$\leq O(n^{۳})$
SE۹۹	$O(n^{۲})$	LTL^T	
GMW–I	$O(n^{۲})$	LTL^T-MS79	$O(n^{۲})$
GMW–II	$O(n^{۲})$	LTL^T-CH98	$O(n^{۲})$
SE–I	$O(n^{۲})$		

الگوریتم های LDL^T تغییریافته

ماتریس زیر تجزیه LDL^T ندارد:

$$A = \begin{bmatrix} ۰ & ۱ \\ ۱ & ۰ \end{bmatrix}$$

ماتریس زیر تجزیه دارد اما تغییراتی که ایجاد می کنیم کران دار نیست.

$$A = \begin{bmatrix} \epsilon & ۱ \\ ۱ & ۰ \end{bmatrix} = \begin{bmatrix} ۱ & ۰ \\ ۱/\epsilon & ۱ \end{bmatrix} \begin{bmatrix} \epsilon & ۰ \\ ۰ & -۱/\epsilon \end{bmatrix} \begin{bmatrix} ۱ & ۱/\epsilon \\ ۰ & ۱ \end{bmatrix}$$

الگوریتم های LDL^T تغییریافته

$$E = \mathrm{diag}\left(\delta_۱, \delta_۲, \ldots, \delta_n\right)$$

مکمل shur در قدم k ام :

$$A_k = \begin{bmatrix} a_k & c_k^T \\ c_k & \bar{A}_k \end{bmatrix}$$

$$L(k+۱:n,k) = \frac{c_k}{a_k+\delta_k}, D(k,k) = a_k + \delta_k \quad (\text{مدل ۱})$$

$$A_{k+۱} = \bar{A}_k - \frac{c_k c_k^T}{a_k + \delta_k}$$

GMW81(Gill et al. (1981))

$$a_k + \delta_k = \max\left\{\delta, |a_k|, \frac{\|c_k\|_\infty^2}{\beta^2}\right\}; \delta = \epsilon_M$$

که نتیجه می‌دهد:

$$\|E\|_2 = o(n^2)$$

SE90, SE99 (Schnabel and Esko (1990, 1999))

تعریف (شعاع Gerschgorin)

شعاع Gerschgorin ماتریس A برابر است با:

$$R_i(A) = \sum_{j \neq i} |a_{i,j}| \text{and} C_i(A) = \{z : \|z - a_{i,i}\| \leq R_i(A)\}$$

قضیه Gerschgorin نشان می‌دهد که مقادیر ویژه در اجتماع دایره های Gerschgorin هستند.(Horn and Johnson, (1985))

SE90, SE99 (Schnabel and Esko (1990, 1999))

- اگر قرار دهیم $\delta_k = \max\{۰, -a_k + \|c_k\|_۱\}$ مقادیر ویژه مثبت خواهند بود.
- این روش ویژگی اول را ندارد. بنابراین یک الگوریتم دو فازی ارائه می‌کنیم.

SE90, SE99 (Schnabel and Esko (1990, 1999))

Algorithm 1 Phase 1 of a 2-Phase Strategy.

{Given a symmetric $A \in R^{n\times n}$ and a small tolerance $\delta > 0$.}
$A_1 := A, k := 1$
Pivot on the maximum diagonal element of A_1.
{Denote $A_k = \begin{bmatrix} a_k & c_k^T \\ c_k & \bar{A}_k \end{bmatrix}$, then $\mathrm{Diag}(\bar{A}_k) \le a_k I_{n-k}$ after pivoting.}
if $a_1 \ge \delta$ **then**
  while $\mathrm{Diag}(\bar{A}_k - \frac{c_k c_k^T}{a_k}) \ge \delta I_{n-k}$ and $k < n$ **do**
    $A_{k+1} := \bar{A}_k - \frac{c_k c_k^T}{a_k}$
    $k := k + 1$
    Pivot on maximum diagonal of A_k.
  end while
end if

فاز یک برای برخی ماتریس ها $\|E\|$ بزرگی دارد. بنابراین فاز یک ریلکس شده را انجام می‌دهیم.

SE90, SE99 (Schnabel and Esko (1990, 1999))

Algorithm 2 Relaxed Phase 1 of a 2-Phase Strategy.

{Given a symmetric $A \in R^{n\times n}$, $\delta > 0$ and $0 < \mu \leq 1$.}
$\eta := \max_{1\leq i\leq n} |A_{ii}|$
if $\mathrm{Diag}(A) \geq -\mu\eta I_n$ **then**
  $A_1 := A$, $k := 1$
  Pivot on the maximum diagonal element of A_1.
  {Denote $A_k = \begin{bmatrix} a_k & c_k^T \\ c_k & \bar{A}_k \end{bmatrix}$, then $\mathrm{Diag}(\bar{A}_k) \leq a_k I_{n-k}$ after pivoting.}
  while $a_k \geq \delta$ and $\mathrm{Diag}(A_k) \geq -\mu a_k I_{n-k+1}$ and $\mathrm{Diag}(\bar{A}_k - \frac{c_k c_k^T}{a_k}) \geq -\mu\eta I_{n-k}$ and $k < n$ **do**
    $A_{k+1} := \bar{A}_k - \frac{c_k c_k^T}{a_k}$
    $k := k + 1$
    Pivot on maximum diagonal of A_k.
  end while
end if

SE90, SE99 (Schnabel and Esko (1990, 1999))

- فاز ۲ :

$$\delta_k = \max\left\{\delta_{k-1}, -a_{k-1} + \max\|c_k\|_1, \delta\right\} \qquad (1)$$

الگوریتم GMW-I

- از فاز اول SE۹۹ (ریکلس شده) و فاز دوم GMW۸۱ استفاده می‌کند و قرار می‌دهد

$$\delta_k = ۰ \quad ; \quad k = ۱, ۲, \ldots, K$$

- خواهیم داشت:

$$\|E\|_۲ \leq o(n)$$

الگوریتم GMW-II

- نسخه نوع دوم از الگوریتم GMW۸۱ :

$$a_k + \delta_k = \max\left\{\delta, a_k + \delta_{k-۱}, \frac{\|c_k\|_\infty^۲}{\beta^۲}\right\}$$

- خواهیم داشت:

$$\|E\|_۲ = O(n^۲)$$

- می‌توانیم الگوریتم ریلکس شده دو فازی را در GMW-II نیز به کار ببریم.
- خواهیم داشت:

$$\|E\|_۲ = O(n)$$

تجزیه LBL^T

- برای ماتریس A متقارن:

$$A = LBL^T$$

که B قطری بلوکی با بلوک های مرتبه ۱ یا ۲ است.

- اگر $\hat{B} = B + \Delta B > ۰$ داریم:

$$P(A + E)P^T = L\hat{B}L^T > ۰$$

روش Moré و Sorensen (MS۷۹)

- برای هر بلوک ۱ × ۱ در B :

$$\hat{d} = \max\{\delta, |d|\}$$

- برای هر بلوک ۲ × ۲ که $D = U \begin{bmatrix} \lambda_۱ & ۰ \\ ۰ & \lambda_۲ \end{bmatrix} U^T$:

$$\hat{\lambda}_i = \max\{\delta, |\lambda_i|\}$$
$$\delta = \epsilon_M$$

روش Cheng و Higham (CH۹۸)

- برای هر بلوک ۱ × ۱ در B :

$$\hat{d} = \max\{\delta, d\}$$

- برای هر بلوک ۲ × ۲ که $D = U \begin{bmatrix} \lambda_۱ & ۰ \\ ۰ & \lambda_۲ \end{bmatrix} U^T$:

$$\hat{\lambda}_i = \max\{\delta, \lambda_i\}$$
$$\delta = \sqrt{u}\|A\|_\infty$$

محور گیری در LBL^T

Table 4 Comparison costs of various pivoting strategies for the LBL^T factorization

Symmetric matrix	General		Tridiagonal	
Case	Worst	Best	Worst	Best
BP	$O(n^3)$		$O(n^2)$	
FBP	$O(n^3)$	$O(n^2)$	$O(n^2)$	$O(n)$
BBK	$O(n^3)$	$O(n^2)$	$O(n^2)$	$O(n)$

تجزیه LTL^T

$$PAP^T = LTL^T$$

- به طوری که T ماتریسی متقارن و سه قطری است و L پایین مثلثی واحد است.
- درایه های غیر قطری L از یک کمتر هستند.
- ستون اول L صفر است.

روش LBL^T–LTL^T ساندویچی

-

$$\hat{T} = T + \Delta T \quad ; \quad \hat{T} > ۰$$

- به طور معادل

$$\hat{A} = P(A + E)P^T = L\hat{T}L^T$$
$$\hat{A} > ۰ \Leftrightarrow \hat{T} > ۰$$

روش LBL^T-LTL^T ساندویچی

- ابتدا تجزیه LTL^T را محاسبه می‌کنیم. $PAP^T = LTL^T$
- یکی از الگوریتم های LDL^T یا LBL^T را روی T اعمال می‌کنیم. در اینصورت خاصیت ۱ برقرار است.
-

قضیه ((2.5) از (Fang, O'Leary (2006))

تجزیه های LDL^T و LBL^T ساختار سه تایی را(triadic) حفظ می‌کنند. یعنی ماتریس های L, D, B ماتریس هایی سه تایی هستند.

روش LBL^T-LTL^T ساندویچی

Table 4 Comparison costs of various pivoting strategies for the LBL^T factorization

Symmetric matrix	General		Tridiagonal	
Case	Worst	Best	Worst	Best
BP	$O(n^3)$		$O(n^2)$	
FBP	$O(n^3)$	$O(n^2)$	$O(n^2)$	$O(n)$
BBK	$O(n^3)$	$O(n^2)$	$O(n^2)$	$O(n)$

- روش LTL^T-MS79 با $\delta = \epsilon_M$
- روش LTL^T-CH98 با $\delta = \sqrt[3]{\epsilon_M^2}$

روش LBL^T-LTL^T ساندویچی

- برای A تجزیه LTL^T را محاسبه کنیم: $PAP^T = LTL^T$
- برای T تجزیه LDL^T را محاسبه کنیم: $\tilde{P}T\tilde{P}^T = \tilde{L}\tilde{B}\tilde{L}^T$
- خواهیم داشت: $PAP^T = L\tilde{P}^T\tilde{L}\tilde{B}\tilde{L}^T\tilde{P}L^T$
- انحراف $\Delta\tilde{B}$ را محاسبه می‌کنیم به طوری که $\tilde{B} + \Delta\tilde{B} > ۰$
- تجزیه $(T + \Delta T)$ برابر است با $\tilde{P}(T + \Delta T)\tilde{P}^T = \tilde{L}(\tilde{B} + \Delta\tilde{B})\tilde{L}^T$
- تجزیه LTL^T تغییر یافته برابر است با:

$$P(A + E)P^T = L\tilde{P}^T\tilde{L}(\tilde{B} + \Delta\tilde{B})\tilde{L}^T\tilde{P}L^T$$

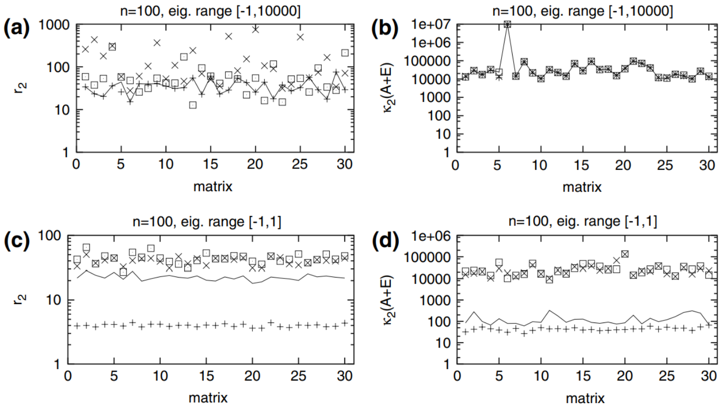
معیار های بزرگی E

- $$r_۲ = \frac{\|E\|_۲}{|\lambda_{min}(A)|}$$
- $$r_F = \frac{\|E\|_F}{\left(\sum_{\lambda_i(A)<۰} \lambda_i(A)^۲\right)^{\frac{۱}{۲}}}$$

Algorithm	r_2	r_F	$\kappa_2(A+E)$
GMW81	2.733	2.674	4.50×10^4
GMW-I	3.014	2.739	4.51×10^4
GMW-II	2.564	2.489	1.64×10^5
SE90	2.78×10^3	3.70×10^3	8.858
SE99	1.759	1.779	1.04×10^{10}
SE-I	3.346	3.289	3.61×10^4
MS79	3.317	2.689	3.33×10^4
CH98	1.659	1.345	9.88×10^7
LTL^T-MS79	3.317	2.689	3.33×10^4
LTL^T-CH98	1.658	1.344	6.74×10^{10}

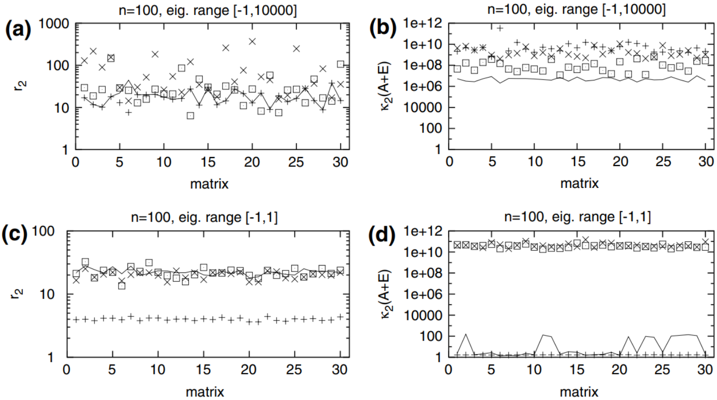
معیار های بزرگی E

الگوریتم های نوع ۱ :



معیار های بزرگی E

الگوریتم های نوع ۲ :



مرجع I

O'Leary P. Dianne · Fang Haw–ren
approaches new with catalog a algorithms: Cholesky Modified