

Um “algoritmo espectral” para redução do envelope de matrizes esparsas

Leonardo D. Secchin

Universidade Federal do Espírito Santo
Mestrado em Informática

Maio de 2009

Seja dada $A \in \mathbb{R}^{n \times n}$ definida positiva (note que A é simétrica com diagonal positiva). A fica completamente determinada por seu triângulo inferior e sua diagonal. Definimos

$$\text{row}(A, i) = \{j; a_{ij} \neq 0, 1 \leq j \leq i\}.$$

Para cada linha i de A , definimos

$$f_i(A) = \min \{j; j \in \text{row}(A, i)\},$$

$$r_i(A) = i - f_i(A).$$

Observe que como $a_{ii} \neq 0, \forall i$, temos $1 \leq f_i(A) \leq i, \forall i$. O *envelope* de A é definido como

$$\text{env}(A) = \{(i, j); f_i(A) \leq j < i, i = 1, \dots, n\}.$$

Denotaremos por $\text{esize}(A) = |\text{env}(A)|$ o tamanho do envelope de A .

A fatoração de Cholesky de A pode ser feita ignorando os termos nulos fora de $\text{env}(A)$. Isto pode ser feito em tempo

$$\frac{1}{2} \sum_{i=2}^n r_i^2(A) + O\left(\sum_{i=2}^n r_i(A)\right).$$

Uma boa medida para isso é

$$\text{ework}(A) = \sum_{i=2}^n r_i^2(A).$$

Os parâmetros $\text{esize}(A)$ e $\text{ework}(A)$ dependem da ordem das linhas/colunas de A . Definimos então

$$\text{esize}_{\min}(A) = \min_{P \text{ perm}} \text{esize}(P^t A P),$$

onde $P \text{ perm}$ significa “ P é matriz de permutação”. Definimos também $\text{ework}_{\min}(A)$ de maneira análoga. Vale observar que esses mínimos em geral não são atingidos para uma mesma matriz de permutação.

Agora, associamos um grafo à matriz A , da maneira que segue. Tome $G(V, E)$ não orientado com $V = \{v_1, \dots, v_n\}$ e $(v_i, v_j) \in E \Leftrightarrow (A - \text{diag}(A))_{ij} \neq 0$. G é o grafo sem laços cuja matriz de adjacência “tem a estrutura de A ”.

O produto P^tAP equivale a uma reordenação dos vértices de G . Podemos então tratar os problemas de minimizar $\text{esize}(P^tAP)$ e $\text{ework}(P^tAP)$ buscando a melhor ordenação dos vértices de G . Dada uma ordenação (bijeção) $\alpha : V \rightarrow \{1, \dots, n\}$ dos vértices de G , definimos

$$r(v, \alpha) = \max \{ \alpha(v) - \alpha(w); w \in \{v\} \cup \text{adj}(v), \alpha(w) \leq \alpha(v) \},$$

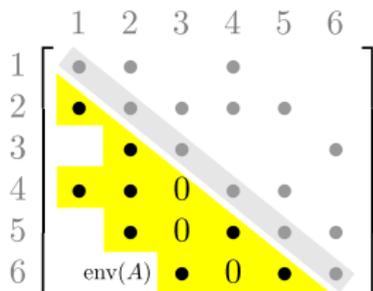
onde $\text{adj}(v)$ são os vértices adjacentes à $v \in V$. Também definimos

$$\text{esize}(G, \alpha) = \sum_{v \in V} r(v, \alpha), \quad \text{ework}(G, \alpha) = \sum_{v \in V} r^2(v, \alpha),$$

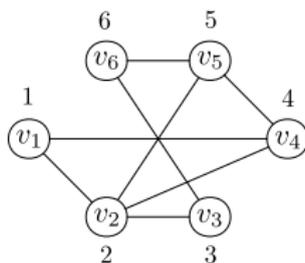
$$\text{esize}_{\min}(G) = \min_{\alpha} \text{esize}(G, \alpha) \quad \text{e} \quad \text{ework}_{\min}(G) = \min_{\alpha} \text{ework}(G, \alpha)$$

Temos $\text{esize}(G, \alpha) = \text{esize}(P^tAP)$, $\text{ework}(G, \alpha) = \text{ework}(P^tAP)$, onde α “=” $P[1 \dots n]^t$.

Para exemplificar, consideramos uma matriz de ordem 6, cuja estrutura de esparsidade é dada abaixo, e seu grafo associado.

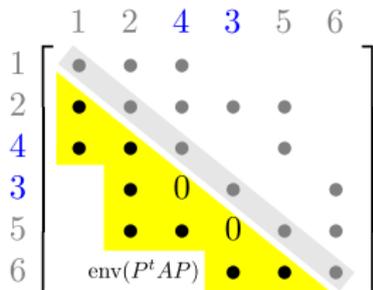


$$\begin{aligned} \text{esize}(A) &= 11 \\ \text{ework}(A) &= 29 \end{aligned}$$

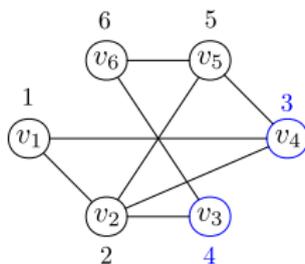


$$\begin{aligned} \text{esize}(G, id) &= 11 \\ \text{ework}(G, id) &= 29 \end{aligned}$$

$r(v_1, id)$	0
$r(v_2, id)$	1
$r(v_3, id)$	1
$r(v_4, id)$	3
$r(v_5, id)$	3
$r(v_6, id)$	3
$\sum r$	11
$\sum r^2$	29



$$\begin{aligned} \text{esize}(P^t A P) &= 10 \\ \text{ework}(P^t A P) &= 22 \end{aligned}$$



$$\begin{aligned} \text{esize}(G, \alpha) &= 10 \\ \text{ework}(G, \alpha) &= 22 \end{aligned}$$

$r(v_1, \alpha)$	0
$r(v_2, \alpha)$	1
$r(v_3, \alpha)$	2
$r(v_4, \alpha)$	2
$r(v_5, \alpha)$	3
$r(v_6, \alpha)$	2
$\sum r$	10
$\sum r^2$	22

onde P troca as linhas 3 e 4, e $(\alpha(v_1), \dots, \alpha(v_6)) = (1, 2, 4, 3, 5, 6)$.

A fim de minimizar $\text{esize}(G, \alpha)$ e $\text{ework}(G, \alpha)$, consideramos o problema de r -soma ($1 \leq r < \infty$), que consiste em minimizar, sobre todas P de permutação,

$$\sigma_r(P^t A P) = \left(\sum_{i=1}^n \sum_{j \in \text{row}(P^t A P, i)} (i-j)^r \right)^{\frac{1}{r}}.$$

Sendo $\sigma_{r \min}(A) = \min_{P \text{ perm}} \sigma_r(P^t A P)$, é possível mostrar que

$$\text{esize}_{\min}(A) \leq \sigma_{1 \min}(A), \quad \text{ework}_{\min}(A) \leq \sigma_{2 \min}^2(A)$$

$$\text{e } \sigma_{1 \min}(A) \leq \sqrt{|E|} \sigma_{2 \min}(A) \leq \sqrt{\text{esize}_{\min}(A)} \sigma_{2 \min}(A).$$

Apresentaremos uma maneira de aproximar $\sigma_{2 \min}^2(A)$ via laplaciana de G . Isso de certa forma é suficiente pois vemos que

$$\text{esize}_{\min}(A) \leq \sigma_{2 \min}^2(A) \quad \text{e} \quad \text{ework}_{\min}(A) \leq \sigma_{2 \min}^2(A).$$

Observamos somente que o problema de 1-soma é mais custoso computacionalmente que o de 2-soma.

Além de aproximar $\sigma_{2\min}^2(A)$, apresentaremos uma maneira de recuperar uma matriz de permutação que aproxima $\sigma_{2\min}(A)$.

Primeiramente, sendo $Q = \text{diag}(d(v_1), \dots, d(v_n)) - \text{adj}G$ a *laplaciana* de G , sejam $\lambda_1(Q) \leq \dots \leq \lambda_n(Q)$ seus autovalores. É possível mostrar que

$$\frac{1}{12} \lambda_2(Q) n(n^2 - 1) \leq \sigma_{2\min}^2(A) \leq \frac{1}{12} \lambda_n(Q) n(n^2 - 1).$$

Sendo $p = [1 \ \dots \ n]^t$, é possível mostrar que

$$\sigma_{2\min}^2(A) = \min_{X \text{ perm}} \text{tr} \, QXBX^t,$$

onde $B = pp^t \in \mathbb{R}^{n \times n}$ tem posto 1 (daí a facilidade computacional em relação à problemas de 1-soma). A fim de aproximar soluções deste problema, observamos que

$$X \text{ é matriz de permut.} \quad \Leftrightarrow \quad Xu = X^t u = u, \quad X^t X = I_n \quad \text{e} \quad X \geq 0,$$

onde $u = \frac{1}{\sqrt{n}}(1, 1, \dots, 1) \in \mathbb{R}^n$ (u é unitário).

Seja $V \in \mathbb{R}^{n \times (n-1)}$ uma base ortornormal do complemento do espaço gerado por u . Temos

$$V^t u = 0 \quad \text{e} \quad P = [u \ V] \in \mathbb{R}^{n \times n} \text{ ortornormal.}$$

Mais ainda,

$$P^t X P = \begin{bmatrix} 1 & 0^t \\ 0 & Y \end{bmatrix}, \quad \text{onde } Y = V^t X V.$$

Daí

$$X = P \begin{bmatrix} 1 & 0^t \\ 0 & Y \end{bmatrix} P^t = uu^t + VYV^t.$$

Com um pouco mais de argumentação o problema $\sigma_{2\min}^2(A) = \min_{X \text{ perm}} \text{tr} QXBX^t$, pode ser escrito como

$$\min \text{tr} \bar{Q} Y \bar{B} Y^t, \quad \text{s.a.} \quad YY^t = I_{n-1}, \quad VYV^t \geq -uu^t,$$

onde $\bar{Q} = V^t Q V$ e \bar{B} análogo. Relaxando a restrição $VYV^t \geq -uu^t$, temos que $Y = RS^t$ é solução ótima, onde R (respect. S) é matriz de autovetores de \bar{Q} (respect. \bar{B}), correspondentes aos autovalores de \bar{Q} (respect. \bar{B}) ordenados.

Daí $X_0 = uu^t + VRS^tV^t$ aproxima $\sigma_{2\min}^2$ e atinge o limitante inferior

$$\sigma_2^2(X_0^tAX_0) = \frac{1}{12}\lambda_2(Q)n(n^2 - 1).$$

Notemos no entanto que não necessariamente $X_0 \geq 0$. A estratégia agora é encontrar uma matriz de permutação Z que melhor aproxima X_0 . Sabemos que quando uma matriz C é definida positiva, existe uma única $C^{\frac{1}{2}}$ definida positiva tal que $C^{\frac{1}{2}}C^{\frac{1}{2}} = C$. Temos $Q + \alpha I_n$ definida positiva para qualquer $\alpha > 0$ pois

$$x^t(Q + \alpha I_n)x = x^tQx + \alpha|x|^2 \geq \alpha|x|^2 > 0 \quad \text{se } \alpha > 0 \text{ e } x \neq 0$$

donde usamos o fato de que

$$x^tQx = \sum_{j \leq i, a_{ij} \neq 0} (x_i - x_j)^2 \geq 0, \quad \forall x \in \mathbb{R}^n \quad (Q \text{ é semidefinida positiva}).$$

Agora, daremos uma nova formulação ao problema de minimizar $\sigma_2^2(Z^tAZ)$ sobre $\{Z; Z \text{ perm}\}$. Da igualdade acima, vemos que

$$p^tQp = \sigma_2^2(A) = \sum_{i=1}^n \sum_{j \in \text{row}(A,i)} (i - j)^2 \quad (p = [1 \cdots n]^t).$$

Observamos que

$$(Zp)^t Q(Zp) = p^t (Z^t QZ) p = \sum_{i=1}^n \sum_{j \in \text{row}(Z^t AZ, i)} (i - j)^2 = \sigma_2^2(Z^t AZ).$$

Também

$$\begin{aligned} \|(Q + \alpha I_n)^{\frac{1}{2}} Zp\|^2 &= \left[(Q + \alpha I_n)^{\frac{1}{2}} Zp \right]^t (Q + \alpha I_n)^{\frac{1}{2}} Zp \\ &= (Zp)^t Q(Zp) + \alpha (Zp)^t I_n (Zp) = (Zp)^t Q(Zp) + \underbrace{\alpha p^t p}_{\text{cte}}, \end{aligned}$$

onde usamos o fato de $Z^t Z = I_n$ e $(Q + \alpha I_n)^{\frac{1}{2}}$ ser simétrica. Portanto o problema de minimizar $\sigma_2^2(Z^t AZ)$ é descrito como

$$\min_{Z \text{ perm}} \|(Q + \alpha I_n)^{\frac{1}{2}} Zp\|^2.$$

Fazendo $Z = X_0 + (Z - X_0)$, chegamos à

$$\begin{aligned} & \| (Q + \alpha I_n)^{\frac{1}{2}} Z p \|^2 = \\ & \| (Q + \alpha I_n)^{\frac{1}{2}} X_0 p \|^2 + 2 \operatorname{tr} p^t (Z - X_0) (Q + \alpha I_n) X_0 p + \| (Q + \alpha I_n)^{\frac{1}{2}} (Z - X_0) p \|^2. \end{aligned}$$

Descartando os termos constantes e o último, Z pode ser aproximado por

$$Z \in \arg \min_{Z \text{ perm}} \operatorname{tr} (Q + \alpha I_n) X_0 B Z^t.$$

Substituindo $X_0 = uu^t + VRS^tV^t$ e fazendo (muitas!) contas chegamos à

$$Z \in \arg \min_{Z \text{ perm}} \operatorname{tr} Z^t x_2 p^t,$$

onde x_2 é autovetor de Q associado à $\lambda_2(Q)$.

A solução do problema anterior é uma matriz Z correspondente à permutação que, ou deixa as coordenadas de x_2 em ordem não-crescente, ou não-decrescente (já que $-x_2$ também é autovetor). Supondo que o grafo induzido G é conexo, temos o algoritmo:

Algoritmo 1

- 1 Dada A definida positiva, tome a matriz Q laplaciana do grafo associado à A ;
 - 2 Calcule um autovetor x_2 de Q associado ao autovalor $\lambda_2(Q)$ (isto pode ser feito pelo algoritmo de Lanczos);
 - 3 Ordene as coordenadas de x_2 em ordem não-crescente e faça a matriz de permutação Z_1 correspondente. Faça o mesmo em ordem não-decrescente e tome a matriz de permutação Z_2 correspondente. Escolha a permutação (Z_1 ou Z_2) que leva ao menor $\text{esize}(Z^tAZ)$.
-

Últimas observações

De acordo com as referências:

- O presente algoritmo mostra-se mais eficiente, em tempo de execução e consumo de memória, que o algoritmo Reverse Cuthill-McKee.
- O desempenho do algoritmo pode ser melhorado se usarmos perturbações na diagonal de Q . Uma perturbação por um vetor \bar{d} é dada por $\bar{Q} = Q + \text{diag}(\bar{d})$.
- O passo 1 requer o cálculo dos graus de todos os vértices de G , que pode ser feito com complexidade quadrática (basta contar em cada linha de A o número de elementos não nulos).
- O algoritmo de Lanczos utilizado para o cálculo do autovetor x_2 (passo 2) é paralelizável.
- O passo 3 requer a ordenação de x_2 , que pode ser feito eficientemente pelo *Quicksort* (complexidade quadrática no pior caso).

Referências

-  Barnard, S. T., Pothen A., Simon H. D. *A spectral algorithm for envelope reduction of sparse matrices*, Numerical Linear Algebra with Applications 2:317-334, 1995.
-  George A., Pothen A. *An analysis of spectral envelope-reduction via quadratic assignment problems*. SIAM Journal of Matrix Analysis and its Applications, 18:706-732, 1997.
-  Saad Y. *Numerical methods for large eigenvalue problems*. Manchester University Press, Oxford, 1992.
-  Coleman T. F. *Large sparse numerical optimization*. Springer-Verlag, 1984.
-  Lima E. L. L. *Álgebra linear*. 5ed. Coleção Metemática Universitária, IMPA, 2001.