Newton’s method in eigenvalue optimization for incomplete pairwise comparison matrices

Kristóf Ábele-Nagy
Eötvös Loránd University (ELTE);
Corvinus University of Budapest (BCE)

Sándor Bozóki
Computer and Automation Research Institute,
Hungarian Academy of Sciences (MTA SZTAKI);
Corvinus University of Budapest (BCE)

15 December, 2010

Outline

- Pairwise comparison matrix
- Incomplete pairwise comparison matrix
- Eigenvalue optimization
- Cyclic coordinates
- Newton’s method in one variable
- Newton’s method in higher dimensions

Given n objects with weights $w_1, w_2, w_3, \dots, w_n$. The **pairwise comparison matrix** is defined as follows:

$$\begin{pmatrix} 1 & \frac{w_1}{w_2} & \frac{w_1}{w_3} & \dots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & 1 & \frac{w_2}{w_3} & \dots & \frac{w_2}{w_n} \\ \frac{w_3}{w_1} & \frac{w_3}{w_2} & 1 & \dots & \frac{w_3}{w_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \frac{w_n}{w_3} & \dots & 1 \end{pmatrix},$$

where

$$\begin{aligned} w_{ij} &> 0, \\ w_{ij} &= \frac{1}{w_{ji}}, \\ w_{ij} &= w_{ik} w_{kj}. \end{aligned}$$

for any i, j, k indices.

In real decision situations, weights are unknown, but pairwise comparisons can be made:

$$\mathbf{A} = \begin{pmatrix} 1 & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & 1 & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & 1 & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & 1 \end{pmatrix},$$

where

$$\begin{aligned} a_{ij} &> 0, \\ a_{ij} &= \frac{1}{a_{ji}}. \end{aligned}$$

for $i, j = 1, \dots, n$. The aim is to determine the weight vector $\mathbf{w} = (w_1, w_2, \dots, w_n) \in \mathbb{R}^n_+$.

In the **Eigenvector Method (**EM**)** the approximation $\mathbf{w}^{EM}$ of $\mathbf{w}$ is defined by

$$\mathbf{A}\mathbf{w}^{EM} = \lambda_{max}\mathbf{w}^{EM},$$

where λ_{max} denotes the maximal eigenvalue, also known as Perron eigenvalue, of $\mathbf{A}$ and $\mathbf{w}^{EM}$ denotes the the right-hand side eigenvector of $\mathbf{A}$ corresponding to λ_{max}.

By Perron's theorem, $\mathbf{w}^{EM}$ is positive and unique up to a scalar multiplication. The most often used normalization is $\sum\limits_{i=1}^{n} w_i^{EM} = 1.$

Saaty defined the inconsistency ratio as $CR = \frac{\frac{\lambda_{max}-n}{n-1}}{RI_n}$, where λ_{max} is the Perron eigenvalue of the complete pairwise comparison matrix given by the decision maker, and RI_n is defined as $\frac{\overline{\lambda_{max}}-n}{n-1}$, where $\overline{\lambda_{max}}$ is an average value of the Perron eigenvalues of randomly generated $n \times n$ pairwise comparison matrices.

It is well known that $\lambda_{max} \geq n$ and equals to n if and only if the matrix is consistent, i.e., the transitivity property holds. It follows from the definition that CR is a positive linear transformation of λ_{max}.

According to Saaty, larger value of CR indicates higher level of inconsistency and the 10%-rule $(CR \leq 0.10)$ separates acceptable matrices from unacceptable ones.

Incomplete pairwise comparison matrix
(= pairwise comparison matrix with missing elements)

$$\mathbf{A} = \begin{pmatrix} 1 & a_{12} & - & \dots & a_{1n} \\ 1/a_{12} & 1 & a_{23} & \dots & - \\ - & 1/a_{23} & 1 & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/a_{1n} & - & 1/a_{3n} & \dots & 1 \end{pmatrix}.$$

Incomplete pairwise comparison matrix
(= pairwise comparison matrix with missing elements)

$$\mathbf{A} = \begin{pmatrix} 1 & a_{12} & x_1 & \dots & a_{1n} \\ 1/a_{12} & 1 & a_{23} & \dots & - \\ 1/x_1 & 1/a_{23} & 1 & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/a_{1n} & - & 1/a_{3n} & \dots & 1 \end{pmatrix}.$$

Incomplete pairwise comparison matrix
(= pairwise comparison matrix with missing elements)

$$\mathbf{A} = \begin{pmatrix} 1 & a_{12} & x_1 & \dots & a_{1n} \\ 1/a_{12} & 1 & a_{23} & \dots & x_d \\ 1/x_1 & 1/a_{23} & 1 & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/a_{1n} & 1/x_d & 1/a_{3n} & \dots & 1 \end{pmatrix},$$

Incomplete pairwise comparison matrix
(= pairwise comparison matrix with missing elements)

$$\mathbf{A} = \begin{pmatrix} 1 & a_{12} & x_1 & \dots & a_{1n} \\ 1/a_{12} & 1 & a_{23} & \dots & x_d \\ 1/x_1 & 1/a_{23} & 1 & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/a_{1n} & 1/x_d & 1/a_{3n} & \dots & 1 \end{pmatrix},$$

where $x_1, x_2, \dots, x_d \in \mathbb{R}_+$.

Based on the idea above, Shiraishi, Obata and Daigo considered the eigenvalue optimization problems as follows.

In case of one missing element, denoted by x, the $\lambda_{max}(\mathbf{A}(x))$ to be minimized:

$$\min_{x>0} \lambda_{max}(\mathbf{A}(x)).$$

In case of more than one missing elements, arranged in vector $\mathbf{x}$, the aim is to solve

$$\min_{\mathbf{x}>\mathbf{0}} \lambda_{max}(\mathbf{A}(\mathbf{x})).$$

Graph representation of a pairwise comparison matrix

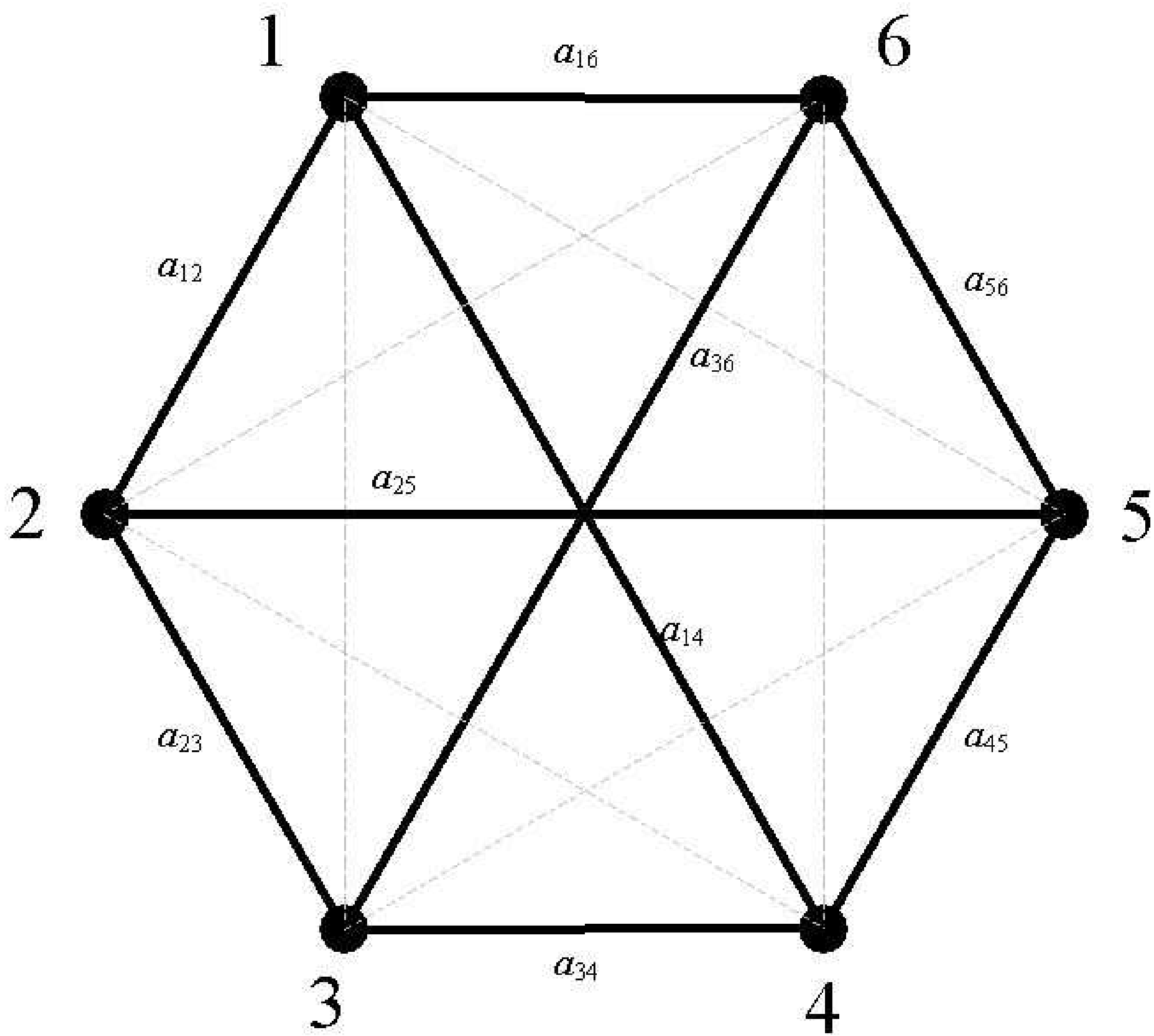
Given $\mathbf{A}$ incomplete pairwise comparison matrix of size $n \times n$. Graph $G = (V, E)$ is defined as follows:

$V = \{1, 2, \ldots, n\}$

$E = \{e(i, j) \,|\, a_{ij}$ (and a_{ji}) are given and $i \neq j\}$

Special case: all the comparisons are given, the corresponding graph is K_n.

Graph representation of a pairwise comparison matrix



	1	2	3	4	5	6
1	-	a_{12}	-	a_{14}	-	a_{16}
2	a_{21}	-	a_{23}	-	a_{25}	-
3	-	a_{32}	-	a_{34}	-	a_{36}
4	a_{41}	-	a_{43}	-	a_{45}	-
5	-	a_{52}	-	a_{54}	-	a_{56}
6	a_{61}	-	a_{63}	-	a_{65}	-

Theorem (B., Fülöp, Rónyai, 2010): The optimal solution of the eigenvalue minimization problem

$$\min_{\mathbf{x}>\mathbf{0}} \lambda_{max}(\mathbf{A}(\mathbf{x})).$$

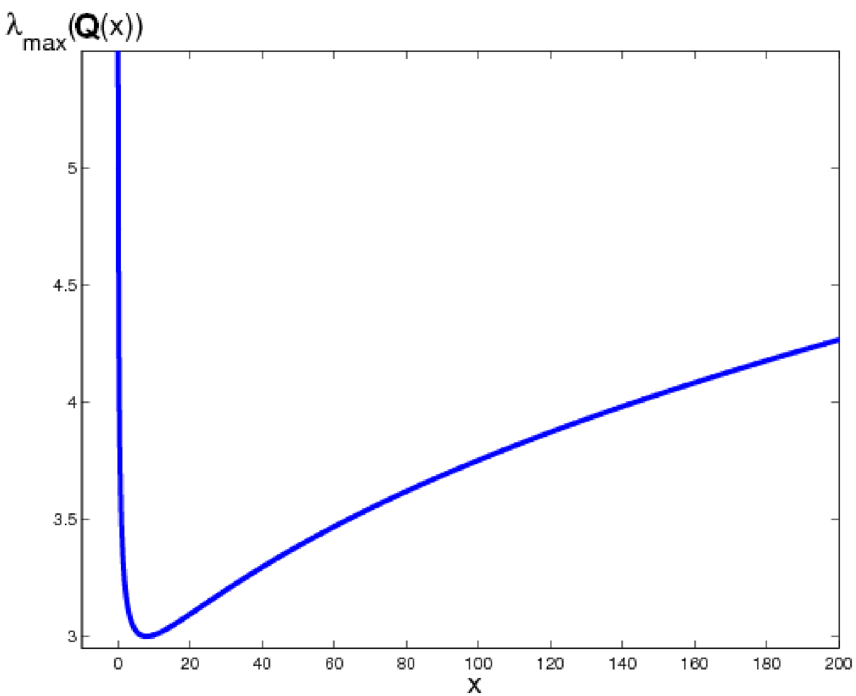
is unique if and only if the graph G corresponding to the incomplete pairwise comparison matrix is *connected*.

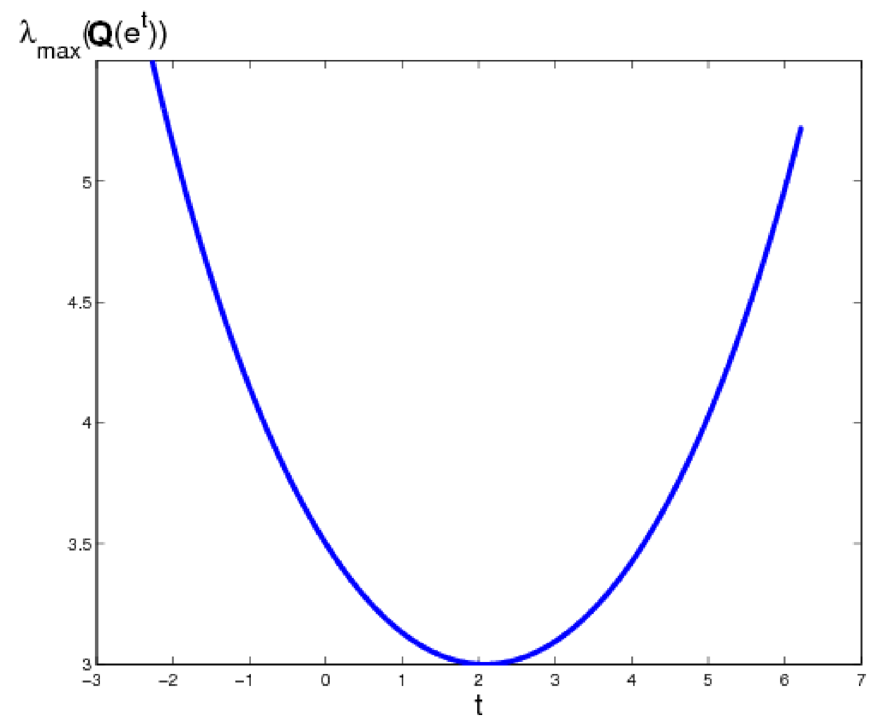
If graph G corresponding to the incomplete pairwise comparison matrix is connected, then by using the exponential parametrization $x_1 = e^{y_1}, x_2 = e^{y_2}, \ldots x_d = e^{y_d}$, the eigenvalue minimization problem is transformed into a strictly convex optimization problem.

Example

$$\mathbf{Q} = \begin{pmatrix} 1 & 2 & x \\ 1/2 & 1 & 4 \\ 1/x & 1/4 & 1 \end{pmatrix}.$$

$\lambda_{max}(\mathbf{Q}(x))$ and, by using the exponential scaling $x = e^t$, $\lambda_{max}(\mathbf{Q}(e^t))$ are plotted.





Algorithms for solving the eigenvalue minimization problem

$$\min_{\mathbf{x}>\mathbf{0}} \lambda_{max}(\mathbf{A}(\mathbf{x})).$$

- cyclic coordinates with Matlab's function *fminbnd*
- cyclic coordinates univariate Newton's method
- multivariate Newton's method

Method of cyclic coordinates

$$\mathbf{M}(\mathbf{x}) = \begin{pmatrix} 1 & 5 & 3 & 7 & 6 & 6 & 1/3 & 1/4 \\ 1/5 & 1 & x_1 & 5 & x_2 & 3 & x_3 & 1/7 \\ 1/3 & 1/x_1 & 1 & x_4 & 3 & x_5 & 6 & x_6 \\ 1/7 & 1/5 & 1/x_4 & 1 & x_7 & 1/4 & x_8 & 1/8 \\ 1/6 & 1/x_2 & 1/3 & 1/x_7 & 1 & x_9 & 1/5 & x_{10} \\ 1/6 & 1/3 & 1/x_5 & 4 & 1/x_9 & 1 & x_{11} & 1/6 \\ 3 & 1/x_3 & 1/6 & 1/x_8 & 5 & 1/x_{11} & 1 & x_{12} \\ 4 & 7 & 1/x_6 & 8 & 1/x_{10} & 6 & 1/x_{12} & 1 \end{pmatrix}$$

Method of cyclic coordinates

Let $x_i^{(k)}$ denote the value of x_i in the k-th step of the iteration, which has d (in the example, $d = 12$) substeps for each k.

For $k = 0$:

Let the initial points be equal to 1 for every variable:

$$x_i^{(0)} := 1 \quad (i = 1, 2, \dots, d).$$

while $\max_{i=1,2,\dots,d} \|x_i^k - x_i^{k-1}\| > T$

$x_i^{(k)} :=$

$\arg\min_{x_i} \lambda_{max}(\mathbf{M}(x_1^{(k)}, \dots, x_{i-1}^{(k)}, x_i, x_{i+1}^{(k-1)}, \dots, x_d^{(k-1)})),\ i = 1, 2, \dots, d$

next k

end while

Method of cyclic coordinates

Focus on $\min_{x_i} \lambda_{max}(\mathbf{M}(x_1^{(k)}, \ldots, x_{i-1}^{(k)}, x_i, x_{i+1}^{(k-1)}, \ldots, x_d^{(k-1)}))$

Matlab's function *fminbnd* solves is directly and fast.

Univariate Newton's method can be also applied.

$$\min_{x>0} \lambda_{max}(\mathbf{A}(x))$$

Let $x = e^t$ and $L(t) = \lambda_{max}(e^t)$.

$$t_{n+1} = t_n - \frac{L'(t_n)}{L''(t_n)} = t_n - \frac{\frac{\partial \lambda_{max}(x)}{\partial x}}{\frac{\partial^2 \lambda_{max}(x)}{(\partial x)^2} \cdot e^{t_n} + \frac{\partial \lambda_{max}(x)}{\partial x}}.$$

By Harker, formal derivatives $\frac{\partial \lambda_{max}(x)}{\partial x}$ and $\frac{\partial^2 \lambda_{max}(x)}{(\partial x)^2}$ are known.

Harker's formula for the first derivative $\frac{\partial \lambda_{max}(x)}{\partial x}$

$$\left(\frac{\partial \lambda_{max}(A)}{\partial a_{ij}} \middle| \, i > j \right) = \left([y(A)_i x(A)_j] - \frac{[y(A)_j x(A)_i]}{[a_{ij}]^2} \right)$$

where vectors $x(A), y(A)$ are the right-hand side and left-hand side eigenvectors of A, respectively. Normalization $y(A)^T x(A) = 1$ is applied.

Harker's formula for the second derivative

$$\frac{\partial^2 \lambda_{max}(A)}{\partial a_{ij} \partial a_{kl}} = (x(A)y(A)^T)_{li} Q^+_{jk} + (x(A)y(A)^T)_{jk} Q^+_{li}$$
$$- \frac{(x(A)y(A)^T)_{ki} Q^+_{jl} + (x(A)y(A)^T)_{jl} Q^+_{ki}}{[a_{kl}]^2}$$
$$- \frac{(x(A)y(A)^T)_{lj} Q^+_{ik} + (x(A)y(A)^T)_{ik} Q^+_{lj}}{[a_{ij}]^2}$$
$$+ \frac{(x(A)y(A)^T)_{kl} Q^+_{il} + (x(A)y(A)^T)_{il} Q^+_{kj}}{[a_{ij}]^2 [a_{kl}]^2}$$

if $i \neq k$ or $j \neq l$,

Harker's formula for the second derivative

$$\frac{\partial^2 \lambda_{max}(A)}{\partial a_{ij} \partial a_{kl}} = \frac{2(x(A)y(A)^T)_{ij}}{[a_{ij}]^3} + 2(x(A)y(A)^T)_{ji} Q^+_{ii}$$

$$-2 \frac{(x(A)y(A)^T)_{ii} Q^+_{jj} + (x(A)y(A)^T)_{jj} Q^+_{ii}}{[a_{ij}]^2}$$

$$+2 \frac{(x(A)y(A)^T)_{ij} Q^+_{ij}}{[a_{ij}]^4}$$

if $i = k$ and $j = l$, where $Q = \lambda_{max}(A) I - A$ and Q^+ denotes the Moore-Penrose pseudoinverse of Q.

Multivariate Newton's method

$$\underline{t}_{n+1} = \underline{t}_n - [HL(\underline{t}_n)]^{-1} \nabla L(\underline{t}_n).$$

Computational results:

- All methods mentioned above are fast enough for typical matrices in multi-attribute decision making;
- in the talk's example (8×8 matrix, 12 variables, 20 cycles with $fminbnd$: 0.3 seconds, 1-variable Newton: 5.3 seconds);
- as a test, in a 150×150 matrix, $\sim$ 10000 variables, 1 cycle takes 1.5 hours with $fminbnd$.

Applications of the results:

- A generalization of the Eigenvector Method for the incomplete case
- CR-inconsistency can be computed during the filling in process, as soon as a connected graph is given
- User may get an automatic warning in case of misprints, detected as a high jump in CR-inconsistency.

Questions:

- How many comparisons are needed, if less than $n(n-1)/2$?
- Thresholds for warning the user?
- Other inconsistency indices, presented by Attila Poesz on Monday (session 2A).

References 1/2

Ábele-Nagy, K. [2010]: Incomplete pairwise comparison matrices in multi-attribute decision making, *Master's Thesis*, Eötvös Loránd University.

Bozóki, S., Fülöp, J., Rónyai, L. [2010]: On optimal completions of incomplete pairwise comparison matrices, *Mathematical and Computer Modelling,* **52**, pp. 318–333.

Harker, P.T. [1987]: Incomplete pairwise comparisons in the analytic hierarchy process. *Mathematical Modelling*, **9**(11), pp. 837–848.

Harker, P.T. [1987]: Derivatives of the Perron root of a positive reciprocal matrix: with application to the Analytic Hierarchy Process. *Applied Mathematics and Computation*, **22**, pp. 217–232.

References 2/2

Saaty, T.L. [1977]: A scaling method for priorities in hierarchical structures, *Journal of Mathematical Psychology,* **15**, pp. 234–281.

Saaty, T.L. [1980]: The analytic hierarchy process, *McGraw-Hill*, New York.

Shiraishi, S., Obata, T., Daigo, M. [1998]: Properties of a positive reciprocal matrix and their application to AHP. *Journal of the Operations Research Society of Japan,* **41**(3), pp. 404–414.

Shiraishi, S., Obata, T. [2002]: On a maximization problem arising from a positive reciprocal matrix in AHP. *Bulletin of Informatics and Cybernetics,* **34**(2), pp. 91–96.

Thank you for attention.

bozoki@sztaki.hu

http://www.sztaki.hu/∼bozoki