Metodi di Newton per MPE

Luca Ferragina

Università di Pisa

21 Dicembre 2016

Piano della presentazione

Definizione di MPE

Definizione

Un' equazione del tipo

$$P(X) = A_0 + A_1 X + \cdots + A_n X^n = 0$$

con $X, A_0, A_1, \ldots, A_n \in \mathbb{R}^{m \times m}$ viene detta *matrix polynomial equation* (MPE).

D'ora in avanti assumeremo su $P(X)$ le seguenti

Ipotesi $(*)$

1. Le matrici $A_n, \ldots, A_2$ e A_0 sono non negative.
2. $-A_1$ è una M-matrice.
3. $(A_2 + \cdots + A_n)\mathbf{1} > 0$.

Definizione di MPE

Definizione

Un' equazione del tipo

$$P(X) = A_0 + A_1 X + \cdots + A_n X^n = 0$$

con $X, A_0, A_1, \ldots, A_n \in \mathbb{R}^{m \times m}$ viene detta *matrix polynomial equation* (MPE).

D'ora in avanti assumeremo su $P(X)$ le seguenti

Ipotesi $(*)$

1. Le matrici $A_n, \ldots, A_2$ e A_0 sono non negative.
2. $-A_1$ è una M-matrice.
3. $(A_2 + \cdots + A_n)\mathbf{1} > 0$.

Definizione di MPE

Definizione

Un' equazione del tipo

$$P(X) = A_0 + A_1 X + \cdots + A_n X^n = 0$$

con $X, A_0, A_1, \ldots, A_n \in \mathbb{R}^{m \times m}$ viene detta *matrix polynomial equation* (MPE).

D'ora in avanti assumeremo su $P(X)$ le seguenti

Ipotesi $(*)$

1. Le matrici $A_n, \ldots, A_2$ e A_0 sono non negative.
2. $-A_1$ è una M-matrice.
3. $(A_2 + \cdots + A_n)\mathbf{1} > 0$.

Connessioni con le catene di Markov

$$P = \begin{bmatrix} B_0 & B_1 & \dots & B_n & B_{n+1} & \dots & \dots \\ \tilde{A}_0 & \tilde{A}_1 & \dots & \tilde{A}_n & 0 & \dots & \dots \\ 0 & \tilde{A}_0 & \tilde{A}_1 & \dots & \tilde{A}_n & 0 & \dots \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

Il calcolo del vettore invariante per una catena di Markov con matrice di iterazione P passa attraverso la risoluzione di un sistema lineare con matrice

$$I - \begin{bmatrix} \tilde{A}_1 & \dots & \tilde{A}_n & 0 & \dots & \dots \\ \tilde{A}_0 & \tilde{A}_1 & \dots & \tilde{A}_n & 0 & \dots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

Connessioni con le catene di Markov

$$P = \begin{bmatrix} B_0 & B_1 & \dots & B_n & B_{n+1} & \dots & \dots \\ \tilde{A}_0 & \tilde{A}_1 & \dots & \tilde{A}_n & 0 & \dots & \dots \\ 0 & \tilde{A}_0 & \tilde{A}_1 & \dots & \tilde{A}_n & 0 & \dots \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

Il calcolo del vettore invariante per una catena di Markov con matrice di iterazione P passa attraverso la risoluzione di un sistema lineare con matrice

$$I - \begin{bmatrix} \tilde{A}_1 & \dots & \tilde{A}_n & 0 & \dots & \dots \\ \tilde{A}_0 & \tilde{A}_1 & \dots & \tilde{A}_n & 0 & \dots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

Connessioni con le catene di Markov

$$P = \begin{bmatrix} B_0 & B_1 & \dots & B_n & B_{n+1} & \dots & \dots \\ \tilde{A}_0 & \tilde{A}_1 & \dots & \tilde{A}_n & 0 & \dots & \dots \\ 0 & \tilde{A}_0 & \tilde{A}_1 & \dots & \tilde{A}_n & 0 & \dots \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

Il calcolo del vettore invariante per una catena di Markov con matrice di iterazione P passa attraverso la risoluzione di un sistema lineare con matrice

$$I - \begin{bmatrix} \tilde{A}_1 & \dots & \tilde{A}_n & 0 & \dots & \dots \\ \tilde{A}_0 & \tilde{A}_1 & \dots & \tilde{A}_n & 0 & \dots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

Grazie alla formula di Wiener-Hopf sappiamo che la matrice precedente si può fattorizzare come

$$\begin{bmatrix} I - A_0^* & -A_1^* & \dots & -A_n^* & 0 & \dots & \dots \\ 0 & I - A_0^* & -A_1^* & \dots & -A_n^* & 0 & \dots \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} trid(-G, I, 0)$$

dove $A_k^* = \sum_{j=k}^{n} \tilde{A}_j G^{j-k}$ e G è la minima soluzione non negativa di $X = \sum_{k=0}^{n} \tilde{A}_k X^k$.
Per cui il problema si riduce a risolvere una MPE con coefficienti

$$A_0 = \tilde{A}_0 \quad A_1 = \tilde{A}_1 - I \quad A_2 = \tilde{A}_2 \quad \dots \quad A_n = \tilde{A}_n$$

Grazie alla formula di Wiener-Hopf sappiamo che la matrice precedente si può fattorizzare come

$$\begin{bmatrix} I - A_0^* & -A_1^* & \dots & -A_n^* & 0 & \dots & \dots \\ 0 & I - A_0^* & -A_1^* & \dots & -A_n^* & 0 & \dots \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} trid(-G, I, 0)$$

dove $A_k^* = \sum_{j=k}^{n} \tilde{A}_j G^{j-k}$ e G è la minima soluzione non negativa di $X = \sum_{k=0}^{n} \tilde{A}_k X^k$.

Per cui il problema si riduce a risolvere una MPE con coefficienti

$$A_0 = \tilde{A}_0 \quad A_1 = \tilde{A}_1 - I \quad A_2 = \tilde{A}_2 \quad \dots \quad A_n = \tilde{A}_n$$

Grazie alla formula di Wiener-Hopf sappiamo che la matrice precedente si può fattorizzare come

$$\begin{bmatrix} I - A_0^* & -A_1^* & \dots & -A_n^* & 0 & \dots & \dots \\ 0 & I - A_0^* & -A_1^* & \dots & -A_n^* & 0 & \dots \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} trid(-G, I, 0)$$

dove $A_k^* = \sum_{j=k}^{n} \tilde{A}_j G^{j-k}$ e G è la minima soluzione non negativa di $X = \sum_{k=0}^{n} \tilde{A}_k X^k$.
Per cui il problema si riduce a risolvere una MPE con coefficienti

$$A_0 = \tilde{A}_0 \quad A_1 = \tilde{A}_1 - I \quad A_2 = \tilde{A}_2 \quad \dots \quad A_n = \tilde{A}_n$$

Il metodo di Newton per la risoluzione di una MPE può essere definito da una matrice di partenza X_0 e dalle iterazioni

$$\begin{cases} \mathcal{D}_{X_i}(H_i) = -P(X_i) \\ X_{i+1} = X_i + H_i \end{cases} \tag{1}$$

dove $\mathcal{D}_X(H) : \mathbb{C}^{m\times m} \mapsto \mathbb{C}^{m\times m}$ è la derivata di Frechet del polinomio $P(X)$ nella direzione H:

$$\mathcal{D}_X(H) = \lim_{t\to 0} \frac{P(X+tH) - P(X)}{t}$$

Il metodo di Newton per la risoluzione di una MPE può essere definito da una matrice di partenza X_0 e dalle iterazioni

$$\begin{cases} \mathcal{D}_{X_i}(H_i) = -P(X_i) \\ X_{i+1} = X_i + H_i \end{cases} \tag{1}$$

dove $\mathcal{D}_X(H) : \mathbb{C}^{m\times m} \mapsto \mathbb{C}^{m\times m}$ è la derivata di Frechet del polinomio $P(X)$ nella direzione H:

$$\mathcal{D}_X(H) = \lim_{t\to 0} \frac{P(X+tH) - P(X)}{t}$$

Il metodo di Newton per la risoluzione di una MPE può essere definito da una matrice di partenza X_0 e dalle iterazioni

$$\begin{cases} \mathcal{D}_{X_i}(H_i) = -P(X_i) \\ X_{i+1} = X_i + H_i \end{cases} \tag{1}$$

dove $\mathcal{D}_X(H) : \mathbb{C}^{m\times m} \mapsto \mathbb{C}^{m\times m}$ è la derivata di Frechet del polinomio $P(X)$ nella direzione H:

$$\mathcal{D}_X(H) = \lim_{t\to 0} \frac{P(X+tH) - P(X)}{t}$$

Esplicitando i conti si può verificare che:

$$\mathcal{D}_X(H) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} H X^{q-1} \Big).$$

La prima equazione di (1) può essere riscritta tramite la funzione *vec* e il prodotto di Kronecker come $D_{X_i} vec(H_i) = vec(-P(X_i))$ dove

$$D_X = \sum_{p=1}^{n} \Big[\sum_{q=1}^{p} (X^{q-1})^T \otimes (A_p X^{p-q}) \Big]$$

Esplicitando i conti si può verificare che:

$$\mathcal{D}_X(H) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} H X^{q-1} \Big).$$

La prima equazione di (1) può essere riscritta tramite la funzione *vec* e il prodotto di Kronecker come $D_{X_i} vec(H_i) = vec(-P(X_i))$ dove

$$D_X = \sum_{p=1}^{n} \Big[\sum_{q=1}^{p} (X^{q-1})^T \otimes (A_p X^{p-q}) \Big]$$

Esplicitando i conti si può verificare che:

$$\mathcal{D}_X(H) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} H X^{q-1} \Big).$$

La prima equazione di (1) può essere riscritta tramite la funzione *vec* e il prodotto di Kronecker come $D_{X_i} vec(H_i) = vec(-P(X_i))$ dove

$$D_X = \sum_{p=1}^{n} \Big[\sum_{q=1}^{p} (X^{q-1})^T \otimes (A_p X^{p-q}) \Big]$$

Convergenza del metodo di Newton

Lemma (1)

Siano $X, Y \in \mathbb{R}^{m \times m}$*, valgono le seguenti disuguaglianze:*

- *Se* $Y > X \geq 0$ *e* $n \geq 2$*, allora:*

$$Y^n + (n-1)X^n - \sum_{q=1}^{n} X^{n-q} Y X^{q-1} > 0$$

- *Se* $Y \geq X \geq 0$ *e* $n \geq 1$*, allora:*

$$Y^n + (n-1)X^n - \sum_{q=1}^{n} X^{n-q} Y X^{q-1} \geq 0$$

Convergenza del metodo di Newton

Lemma (1)

Siano $X, Y \in \mathbb{R}^{m \times m}$, valgono le seguenti disuguaglianze:

- *Se $Y > X \geq 0$ e $n \geq 2$, allora:*

$$Y^n + (n-1)X^n - \sum_{q=1}^{n} X^{n-q} Y X^{q-1} > 0$$

- *Se $Y \geq X \geq 0$ e $n \geq 1$, allora:*

$$Y^n + (n-1)X^n - \sum_{q=1}^{n} X^{n-q} Y X^{q-1} \geq 0$$

Convergenza del metodo di Newton

Lemma (1)

Siano $X, Y \in \mathbb{R}^{m \times m}$, valgono le seguenti disuguaglianze:

- *Se $Y > X \geq 0$ e $n \geq 2$, allora:*

$$Y^n + (n-1)X^n - \sum_{q=1}^{n} X^{n-q} Y X^{q-1} > 0$$

- *Se $Y \geq X \geq 0$ e $n \geq 1$, allora:*

$$Y^n + (n-1)X^n - \sum_{q=1}^{n} X^{n-q} Y X^{q-1} \geq 0$$

Convergenza del metodo di Newton

Lemma (1)

Siano $X, Y \in \mathbb{R}^{m\times m}$*, valgono le seguenti disuguaglianze:*

- *Se* $Y > X \geq 0$ *e* $n \geq 2$*, allora:*

$$Y^n + (n-1)X^n - \sum_{q=1}^{n} X^{n-q} Y X^{q-1} > 0$$

- *Se* $Y \geq X \geq 0$ *e* $n \geq 1$*, allora:*

$$Y^n + (n-1)X^n - \sum_{q=1}^{n} X^{n-q} Y X^{q-1} \geq 0$$

Dimostrazione.

Per $n = 2$ si ha che

$$Y^2 + X^2 - XY - YX = Y(Y - X) - X(Y - X) = (Y - X)^2 > 0.$$

Suppongo l'enunciato vero per n, allora valgono:

$$\left[Y^n + (n-1)X^n - \sum_{q=1}^{n} X^{n-q} Y X^{q-1}\right] X =$$

$$= Y^n X + (n-1)X^{n+1} - \sum_{q=1}^{n} X^{n-q} Y X^q \geq 0$$

$$(Y^n - X^n)(Y - X) = Y^{n+1} - X^n Y - Y^n X + X^{n+1} > 0.$$

Adesso sommiamo queste due relazioni.

Dimostrazione.

Per $n = 2$ si ha che

$$Y^2 + X^2 - XY - YX = Y(Y - X) - X(Y - X) = (Y - X)^2 > 0.$$

Suppongo l'enunciato vero per n, allora valgono:

$$\left[Y^n + (n-1)X^n - \sum_{q=1}^{n} X^{n-q} Y X^{q-1}\right] X =$$

$$= Y^n X + (n-1)X^{n+1} - \sum_{q=1}^{n} X^{n-q} Y X^q \geq 0$$

$$(Y^n - X^n)(Y - X) = Y^{n+1} - X^n Y - Y^n X + X^{n+1} > 0.$$

Adesso sommiamo queste due relazioni.

Dimostrazione.

Per $n = 2$ si ha che

$$Y^2 + X^2 - XY - YX = Y(Y - X) - X(Y - X) = (Y - X)^2 > 0.$$

Suppongo l'enunciato vero per n, allora valgono:

$$\left[Y^n + (n-1)X^n - \sum_{q=1}^{n} X^{n-q} Y X^{q-1}\right] X =$$

$$= Y^n X + (n-1)X^{n+1} - \sum_{q=1}^{n} X^{n-q} Y X^q \geq 0$$

$$(Y^n - X^n)(Y - X) = Y^{n+1} - X^n Y - Y^n X + X^{n+1} > 0.$$

Adesso sommiamo queste due relazioni.

Dimostrazione.

Per $n = 2$ si ha che

$$Y^2 + X^2 - XY - YX = Y(Y - X) - X(Y - X) = (Y - X)^2 > 0.$$

Suppongo l'enunciato vero per n, allora valgono:

$$\left[Y^n + (n-1)X^n - \sum_{q=1}^{n} X^{n-q} Y X^{q-1}\right] X =$$

$$= Y^n X + (n-1)X^{n+1} - \sum_{q=1}^{n} X^{n-q} Y X^{q} \geq 0$$

$$(Y^n - X^n)(Y - X) = Y^{n+1} - X^n Y - Y^n X + X^{n+1} > 0.$$

Adesso sommiamo queste due relazioni.

Dimostrazione.

Per $n = 2$ si ha che

$$Y^2 + X^2 - XY - YX = Y(Y - X) - X(Y - X) = (Y - X)^2 > 0.$$

Suppongo l'enunciato vero per n, allora valgono:

$$\left[Y^n + (n-1)X^n - \sum_{q=1}^{n} X^{n-q} Y X^{q-1}\right] X =$$

$$= Y^n X + (n-1)X^{n+1} - \sum_{q=1}^{n} X^{n-q} Y X^q \geq 0$$

$$(Y^n - X^n)(Y - X) = Y^{n+1} - X^n Y - Y^n X + X^{n+1} > 0.$$

Adesso sommiamo queste due relazioni.

Dimostrazione.

Per $n = 2$ si ha che

$$Y^2 + X^2 - XY - YX = Y(Y - X) - X(Y - X) = (Y - X)^2 > 0.$$

Suppongo l'enunciato vero per n, allora valgono:

$$\left[Y^n + (n-1)X^n - \sum_{q=1}^{n} X^{n-q} Y X^{q-1}\right] X =$$

$$= Y^n X + (n-1)X^{n+1} - \sum_{q=1}^{n} X^{n-q} Y X^q \geq 0$$

$$(Y^n - X^n)(Y - X) = Y^{n+1} - X^n Y - Y^n X + X^{n+1} > 0.$$

Adesso sommiamo queste due relazioni.

Dimostrazione.

$$0 < Y^{n+1} - X^n Y + nX^{n+1} - \sum_{q=1}^{n} X^{n-q} Y X^q$$

$$= Y^{n+1} + nX^{n+1} - \sum_{q=0}^{n} X^{n-q} Y X^q$$

$$= Y^{n+1} + nX^{n+1} - \sum_{q=1}^{n+1} X^{n+1-q} Y X^{q-1}$$

Dimostrazione.

$$
\begin{aligned}
0 &< Y^{n+1} - X^n Y + nX^{n+1} - \sum_{q=1}^{n} X^{n-q} Y X^q \\
&= Y^{n+1} + nX^{n+1} - \sum_{q=0}^{n} X^{n-q} Y X^q \\
&= Y^{n+1} + nX^{n+1} - \sum_{q=1}^{n+1} X^{n+1-q} Y X^{q-1}
\end{aligned}
$$

Dimostrazione.

$$\begin{aligned}
0 &< Y^{n+1} - X^n Y + nX^{n+1} - \sum_{q=1}^{n} X^{n-q} Y X^q \\
&= Y^{n+1} + nX^{n+1} - \sum_{q=0}^{n} X^{n-q} Y X^q \\
&= Y^{n+1} + nX^{n+1} - \sum_{q=1}^{n+1} X^{n+1-q} Y X^{q-1}
\end{aligned}$$

$\square$

Lemma (2)

Siano $X, Y \in \mathbb{R}^{m\times m}$, allora:

- *Se $Y \geq X \geq 0$, allora $\mathcal{D}_Y(H) \geq \mathcal{D}_X(H)$ con $H \geq 0$.*
- *$-D_X$ è una Z-matrice per ogni X matrice non negativa.*

Corollario (1)

Se esiste una Y matrice non negativa tale che $-D_Y$ sia una M-matrice singolare e $Y \geq X \geq 0$ allora $-D_X$ è una M-matrice non singolare.

Lemma (2)

Siano $X, Y \in \mathbb{R}^{m\times m}$, allora:

- *Se $Y \geq X \geq 0$, allora $\mathcal{D}_Y(H) \geq \mathcal{D}_X(H)$ con $H \geq 0$.*
- *$-D_X$ è una Z-matrice per ogni X matrice non negativa.*

Corollario (1)

Se esiste una Y matrice non negativa tale che $-D_Y$ sia una M-matrice singolare e $Y \geq X \geq 0$ allora $-D_X$ è una M-matrice non singolare.

Lemma (2)

Siano $X, Y \in \mathbb{R}^{m \times m}$, allora:

- *Se $Y \geq X \geq 0$, allora $\mathcal{D}_Y(H) \geq \mathcal{D}_X(H)$ con $H \geq 0$.*
- *$-D_X$ è una Z-matrice per ogni X matrice non negativa.*

Corollario (1)

Se esiste una Y matrice non negativa tale che $-D_Y$ sia una M-matrice singolare e $Y \geq X \geq 0$ allora $-D_X$ è una M-matrice non singolare.

Lemma (2)

Siano $X, Y \in \mathbb{R}^{m \times m}$, allora:

- *Se $Y \geq X \geq 0$, allora $\mathcal{D}_Y(H) \geq \mathcal{D}_X(H)$ con $H \geq 0$.*
- *$-D_X$ è una Z-matrice per ogni X matrice non negativa.*

Corollario (1)

Se esiste una Y matrice non negativa tale che $-D_Y$ sia una M-matrice singolare e $Y \geq X \geq 0$ allora $-D_X$ è una M-matrice non singolare.

Dimostrazione.

$$\begin{aligned}
\mathcal{D}_Y(H) - \mathcal{D}_X(H) &= \sum_{p=1}^{n} A_p\Big(\sum_{q=1}^{p} Y^{p-q} H Y^{q-1}\Big) \\
&\quad - \sum_{p=1}^{n} A_p\Big(\sum_{q=1}^{p} X^{p-q} H X^{q-1}\Big) \\
&= \sum_{p=2}^{n} A_p\Big(\sum_{q=1}^{p} Y^{p-q} H Y^{q-1}\Big) \\
&\quad - \sum_{p=2}^{n} A_p\Big(\sum_{q=1}^{p} X^{p-q} H X^{q-1}\Big) \\
&\quad - \sum_{p=2}^{n} A_p\Big(\sum_{q=1}^{p} Y^{p-q} H X^{q-1} - \sum_{q=1}^{p} Y^{p-q} H X^{q-1}\Big)
\end{aligned}$$

Dimostrazione.

$$
\begin{aligned}
\mathcal{D}_Y(H) - \mathcal{D}_X(H) &= \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} Y^{p-q} H Y^{q-1} \Big) \\
&\quad - \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} H X^{q-1} \Big) \\
&= \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} Y^{p-q} H Y^{q-1} \Big) \\
&\quad - \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} H X^{q-1} \Big) \\
&\quad - \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} Y^{p-q} H X^{q-1} - \sum_{q=1}^{p} Y^{p-q} H X^{q-1} \Big)
\end{aligned}
$$

Dimostrazione.

$$
\begin{aligned}
\mathcal{D}_Y(H) - \mathcal{D}_X(H) &= \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} Y^{p-q} H Y^{q-1} \Big) \\
&- \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} H X^{q-1} \Big) \\
&= \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} Y^{p-q} H Y^{q-1} \Big) \\
&- \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} H X^{q-1} \Big) \\
&- \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} Y^{p-q} H X^{q-1} - \sum_{q=1}^{p} Y^{p-q} H X^{q-1} \Big)
\end{aligned}
$$

Dimostrazione.

$$\mathcal{D}_Y(H) - \mathcal{D}_X(H) = \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} Y^{p-q} H (Y^{q-1} - X^{q-1}) \Big) + \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} (Y^{p-q} - X^{p-q}) H Y^{q-1} \Big) \geq 0.$$

Poiché $-A_1$ è una M-matrice non singolare, anche $-(I \otimes A_1)$ lo è.

$$-D_X = -\sum_{p=2}^{n} \Big[\sum_{q=1}^{p} (X^{q-1})^T \otimes (A_p X^{p-q}) \Big] - I \otimes A_1 \leq -I \otimes A_1 = -D_0.$$

Perciò $-D_X$ ha elementi non positivi fuori dalla diagonale. □

Dimostrazione.

$$\mathcal{D}_Y(H) - \mathcal{D}_X(H) = \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} Y^{p-q} H (Y^{q-1} - X^{q-1}) \Big)$$
$$+ \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} (Y^{p-q} - X^{p-q}) H Y^{q-1} \Big) \geq 0.$$

Poiché $-A_1$ è una M-matrice non singolare, anche $-(I \otimes A_1)$ lo è.

$$-D_X = -\sum_{p=2}^{n} \Big[\sum_{q=1}^{p} (X^{q-1})^T \otimes (A_p X^{p-q}) \Big] - I \otimes A_1$$
$$\leq -I \otimes A_1 = -D_0.$$

Perciò $-D_X$ ha elementi non positivi fuori dalla diagonale. □

Dimostrazione.

$$\mathcal{D}_Y(H) - \mathcal{D}_X(H) = \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} Y^{p-q} H (Y^{q-1} - X^{q-1}) \Big) + \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} (Y^{p-q} - X^{p-q}) H Y^{q-1} \Big) \geq 0.$$

Poiché $-A_1$ è una M-matrice non singolare, anche $-(I \otimes A_1)$ lo è.

$$-D_X = -\sum_{p=2}^{n} \Big[\sum_{q=1}^{p} (X^{q-1})^T \otimes (A_p X^{p-q}) \Big] - I \otimes A_1 \leq -I \otimes A_1 = -D_0.$$

Perciò $-D_X$ ha elementi non positivi fuori dalla diagonale.

Dimostrazione.

$$\mathcal{D}_Y(H) - \mathcal{D}_X(H) = \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} Y^{p-q} H (Y^{q-1} - X^{q-1}) \Big) + \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} (Y^{p-q} - X^{p-q}) H Y^{q-1} \Big) \geq 0.$$

Poiché $-A_1$ è una M-matrice non singolare, anche $-(I \otimes A_1)$ lo è.

$$-D_X = -\sum_{p=2}^{n} \Big[\sum_{q=1}^{p} (X^{q-1})^T \otimes (A_p X^{p-q}) \Big] - I \otimes A_1 \leq -I \otimes A_1 = -D_0.$$

Perciò $-D_X$ ha elementi non positivi fuori dalla diagonale. □

Dimostrazione.

$$\mathcal{D}_Y(H) - \mathcal{D}_X(H) = \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} Y^{p-q} H (Y^{q-1} - X^{q-1}) \Big)$$
$$+ \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} (Y^{p-q} - X^{p-q}) H Y^{q-1} \Big) \geq 0.$$

Poiché $-A_1$ è una M-matrice non singolare, anche $-(I \otimes A_1)$ lo è.

$$-D_X = -\sum_{p=2}^{n} \Big[\sum_{q=1}^{p} (X^{q-1})^T \otimes (A_p X^{p-q}) \Big] - I \otimes A_1$$
$$\leq -I \otimes A_1 = -D_0.$$

Perciò $-D_X$ ha elementi non positivi fuori dalla diagonale.

Dimostrazione.

$$\mathcal{D}_Y(H) - \mathcal{D}_X(H) = \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} Y^{p-q} H (Y^{q-1} - X^{q-1}) \Big) + \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} (Y^{p-q} - X^{p-q}) H Y^{q-1} \Big) \geq 0.$$

Poiché $-A_1$ è una M-matrice non singolare, anche $-(I \otimes A_1)$ lo è.

$$-D_X = -\sum_{p=2}^{n} \Big[\sum_{q=1}^{p} (X^{q-1})^T \otimes (A_p X^{p-q}) \Big] - I \otimes A_1 \leq -I \otimes A_1 = -D_0.$$

Perciò $-D_X$ ha elementi non positivi fuori dalla diagonale.

Dimostrazione.

$$\mathcal{D}_Y(H) - \mathcal{D}_X(H) = \sum_{p=2}^{n} A_p\Big(\sum_{q=1}^{p} Y^{p-q}H(Y^{q-1} - X^{q-1})\Big)$$
$$+ \sum_{p=2}^{n} A_p\Big(\sum_{q=1}^{p}(Y^{p-q} - X^{p-q})HY^{q-1}\Big) \geq 0.$$

Poiché $-A_1$ è una M-matrice non singolare, anche $-(I \otimes A_1)$ lo è.

$$-D_X = -\sum_{p=2}^{n}\Big[\sum_{q=1}^{p}(X^{q-1})^T \otimes (A_p X^{p-q})\Big] - I \otimes A_1$$
$$\leq -I \otimes A_1 = -D_0.$$

Perciò $-D_X$ ha elementi non positivi fuori dalla diagonale. $\square$

Lemma (3)

Se esiste una Y matrice positiva tale che $P(Y) \leq 0$, allora:

- *Se $P(X) \geq 0$ e $Y > X \geq 0$, allora $-D_X$ è una M-matrice non singolare.*
- *Se $-D_X$ è una M-matrice non singolare, allora $Y > X^+$ dove $X^+ = X + H$ e H è tale che*

$$\mathcal{D}_X(H) = -P(X)$$

Corollario (2)

Se esiste una Y matrice non negativa tale che $P(Y) \leq 0$, allora, per ogni X tale che $Y \geq X \geq 0$ e $P(X) \geq 0$, $-D_X$ è una M-matrice non singolare e $Y \geq X^+$.

Lemma (3)

Se esiste una Y matrice positiva tale che $P(Y) \leq 0$, allora:

- *Se $P(X) \geq 0$ e $Y > X \geq 0$, allora $-D_X$ è una M-matrice non singolare.*
- *Se $-D_X$ è una M-matrice non singolare, allora $Y > X^+$ dove $X^+ = X + H$ e H è tale che*

$$\mathcal{D}_X(H) = -P(X)$$

Corollario (2)

Se esiste una Y matrice non negativa tale che $P(Y) \leq 0$, allora, per ogni X tale che $Y \geq X \geq 0$ e $P(X) \geq 0$, $-D_X$ è una M-matrice non singolare e $Y \geq X^+$.

Lemma (3)

Se esiste una Y matrice positiva tale che $P(Y) \leq 0$, allora:

- *Se $P(X) \geq 0$ e $Y > X \geq 0$, allora $-D_X$ è una M-matrice non singolare.*
- *Se $-D_X$ è una M-matrice non singolare, allora $Y > X^+$ dove $X^+ = X + H$ e H è tale che*

$$\mathcal{D}_X(H) = -P(X)$$

Corollario (2)

Se esiste una Y matrice non negativa tale che $P(Y) \leq 0$, allora, per ogni X tale che $Y \geq X \geq 0$ e $P(X) \geq 0$, $-D_X$ è una M-matrice non singolare e $Y \geq X^+$.

Lemma (3)

Se esiste una Y matrice positiva tale che $P(Y) \leq 0$, allora:

- *Se $P(X) \geq 0$ e $Y > X \geq 0$, allora $-D_X$ è una M-matrice non singolare.*
- *Se $-D_X$ è una M-matrice non singolare, allora $Y > X^+$ dove $X^+ = X + H$ e H è tale che*

$$\mathcal{D}_X(H) = -P(X)$$

Corollario (2)

Se esiste una Y matrice non negativa tale che $P(Y) \leq 0$, allora, per ogni X tale che $Y \geq X \geq 0$ e $P(X) \geq 0$, $-D_X$ è una M-matrice non singolare e $Y \geq X^+$.

Osservazione

$P(Y) = \sum_{p=0}^{n} A_p Y^p \leq 0 \Rightarrow A_1 Y + A_0 \leq -\sum_{p=2}^{n} A_p Y^p.$

Dimostrazione.

$$
\begin{aligned}
\mathcal{D}_X(Y - X) = \mathcal{D}_X(Y) - \mathcal{D}_X(X) \\
&= \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) - \sum_{p=1}^{n} p A_p X^p \\
&= \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) \\
&- \sum_{p=2}^{n} (p-1) A_p X^p - \sum_{p=1}^{n} A_p X^p
\end{aligned}
$$

Osservazione

$P(Y) = \sum_{p=0}^{n} A_p Y^p \leq 0 \Rightarrow A_1 Y + A_0 \leq -\sum_{p=2}^{n} A_p Y^p.$

Dimostrazione.

$$\begin{aligned}
\mathcal{D}_X(Y - X) &= \mathcal{D}_X(Y) - \mathcal{D}_X(X) \\
&= \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) - \sum_{p=1}^{n} p A_p X^p \\
&= \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) \\
&\quad - \sum_{p=2}^{n} (p-1) A_p X^p - \sum_{p=1}^{n} A_p X^p
\end{aligned}$$

Osservazione

$P(Y) = \sum_{p=0}^{n} A_p Y^p \leq 0 \Rightarrow A_1 Y + A_0 \leq -\sum_{p=2}^{n} A_p Y^p.$

Dimostrazione.

$$\begin{aligned}
\mathcal{D}_X(Y - X) &= \mathcal{D}_X(Y) - \mathcal{D}_X(X) \\
&= \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) - \sum_{p=1}^{n} p A_p X^p \\
&= \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) \\
&\quad - \sum_{p=2}^{n} (p-1) A_p X^p - \sum_{p=1}^{n} A_p X^p
\end{aligned}$$

Osservazione

$P(Y) = \sum_{p=0}^{n} A_p Y^p \leq 0 \Rightarrow A_1 Y + A_0 \leq -\sum_{p=2}^{n} A_p Y^p.$

Dimostrazione.

$$\begin{aligned}
\mathcal{D}_X(Y - X) &= \mathcal{D}_X(Y) - \mathcal{D}_X(X) \\
&= \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) - \sum_{p=1}^{n} p A_p X^p \\
&= \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) \\
&\quad - \sum_{p=2}^{n} (p-1) A_p X^p - \sum_{p=1}^{n} A_p X^p
\end{aligned}$$

Osservazione

$P(Y) = \sum_{p=0}^{n} A_p Y^p \leq 0 \Rightarrow A_1 Y + A_0 \leq -\sum_{p=2}^{n} A_p Y^p.$

Dimostrazione.

$$\begin{aligned}
\mathcal{D}_X(Y - X) &= \mathcal{D}_X(Y) - \mathcal{D}_X(X) \\
&= \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) - \sum_{p=1}^{n} p A_p X^p \\
&= \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) \\
&- \sum_{p=2}^{n} (p-1) A_p X^p - \sum_{p=1}^{n} A_p X^p
\end{aligned}$$

Dimostrazione.

$$
\begin{aligned}
\mathcal{D}_X(Y-X) &= \sum_{p=2}^{n} A_p\Big(\sum_{q=1}^{p} X^{p-q}YX^{q-1}\Big) + A_1Y \\
&- \sum_{p=2}^{n}(p-1)A_pX^p - P(X) + A_0 \\
&\leq \sum_{p=2}^{n} A_p(\sum_{q=1}^{p} X^{p-q}YX^{q-1} - (p-1)X^p) + A_1Y + A_0 \\
&\leq \sum_{p=2}^{n} A_p(\sum_{q=1}^{p} X^{p-q}YX^{q-1} - (p-1)X^p - A_pY^p) \\
&= \sum_{p=2}^{n} A_pN_p < 0
\end{aligned}
$$

Dimostrazione.

$$
\begin{aligned}
\mathcal{D}_X(Y-X) = {} & \sum_{p=2}^{n} A_p\Big(\sum_{q=1}^{p} X^{p-q}YX^{q-1}\Big) + A_1 Y \\
& - \sum_{p=2}^{n}(p-1)A_pX^p - P(X) + A_0 \\
& \leq \sum_{p=2}^{n} A_p(\sum_{q=1}^{p} X^{p-q}YX^{q-1} - (p-1)X^p) + A_1Y + A_0 \\
& \leq \sum_{p=2}^{n} A_p(\sum_{q=1}^{p} X^{p-q}YX^{q-1} - (p-1)X^p - A_pY^p) \\
& = \sum_{p=2}^{n} A_pN_p \leq 0
\end{aligned}
$$

Dimostrazione.

$$
\begin{aligned}
\mathcal{D}_X(Y-X) &= \sum_{p=2}^{n} A_p\Big(\sum_{q=1}^{p} X^{p-q}YX^{q-1}\Big) + A_1 Y \\
&\quad - \sum_{p=2}^{n}(p-1)A_pX^p - P(X) + A_0 \\
&\leq \sum_{p=2}^{n} A_p(\sum_{q=1}^{p} X^{p-q}YX^{q-1} - (p-1)X^p) + A_1Y + A_0 \\
&\leq \sum_{p=2}^{n} A_p(\sum_{q=1}^{p} X^{p-q}YX^{q-1} - (p-1)X^p - A_pY^p) \\
&= \sum_{p=2}^{n} A_pN_p < 0
\end{aligned}
$$

Dimostrazione.

$$\mathcal{D}_X(Y-X) = \sum_{p=2}^{n} A_p\Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1}\Big) + A_1 Y$$
$$- \sum_{p=2}^{n} (p-1) A_p X^p - P(X) + A_0$$
$$\leq \sum_{p=2}^{n} A_p\big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} - (p-1)X^p\big) + A_1 Y + A_0$$
$$\leq \sum_{p=2}^{n} A_p\big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} - (p-1)X^p - A_p Y^p\big)$$
$$= \sum_{p=2}^{n} A_p N_p < 0$$

Dimostrazione.

$$
\begin{aligned}
\mathcal{D}_X(Y-X) &= \sum_{p=2}^{n} A_p\Big(\sum_{q=1}^{p} X^{p-q}YX^{q-1}\Big) + A_1Y \\
&\quad - \sum_{p=2}^{n}(p-1)A_pX^p - P(X) + A_0 \\
&\leq \sum_{p=2}^{n} A_p\big(\sum_{q=1}^{p} X^{p-q}YX^{q-1} - (p-1)X^p\big) + A_1Y + A_0 \\
&\leq \sum_{p=2}^{n} A_p\big(\sum_{q=1}^{p} X^{p-q}YX^{q-1} - (p-1)X^p - A_pY^p\big) \\
&= \sum_{p=2}^{n} A_pN_p < 0
\end{aligned}
$$

Dimostrazione.

Dove $N_p = \sum_{q=1}^{p} X^{p-q}YX^{q-1} - (p-1)X^p - Y^p < 0$ per il lemma (1), $A_p \geq 0$ e vale il minore stretto grazie a $(*)$.

$$\begin{aligned} \mathcal{D}_X(Y - X^+) &= \mathcal{D}_X(Y) - \mathcal{D}_X(X) + P(X) \\ &= \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q}YX^{q-1} \Big) + A_1 Y \\ &- \sum_{p=2}^{n} (p-1)A_p X^p + A_0 \\ &\leq \sum_{p=2}^{n} A_p N_p < 0 \end{aligned}$$

Poiché $-D_X$ è una M-matrice si ha $Y - X^+ > 0$. □

Dimostrazione.

Dove $N_p = \sum_{q=1}^{p} X^{p-q} Y X^{q-1} - (p-1)X^p - Y^p < 0$ per il lemma (1), $A_p \geq 0$ e vale il minore stretto grazie a $(*)$.

$$\begin{aligned}
\mathcal{D}_X(Y - X^+) &= \mathcal{D}_X(Y) - \mathcal{D}_X(X) + P(X) \\
&= \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) + A_1 Y \\
&- \sum_{p=2}^{n} (p-1) A_p X^p + A_0 \\
&\leq \sum_{p=2}^{n} A_p N_p < 0
\end{aligned}$$

Poiché $-D_X$ è una M-matrice si ha $Y - X^+ > 0$.

Dimostrazione.

Dove $N_p = \sum_{q=1}^{p} X^{p-q} Y X^{q-1} - (p-1)X^p - Y^p < 0$ per il lemma (1), $A_p \geq 0$ e vale il minore stretto grazie a $(*)$.

$$\begin{aligned}\mathcal{D}_X(Y - X^+) &= \mathcal{D}_X(Y) - \mathcal{D}_X(X) + P(X) \\ &= \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) + A_1 Y \\ &- \sum_{p=2}^{n} (p-1) A_p X^p + A_0 \\ &\leq \sum_{p=2}^{n} A_p N_p < 0\end{aligned}$$

Poiché $-D_X$ è una M-matrice si ha $Y - X^+ > 0$. □

Dimostrazione.

Dove $N_p = \sum_{q=1}^{p} X^{p-q} Y X^{q-1} - (p-1)X^p - Y^p < 0$ per il lemma (1), $A_p \geq 0$ e vale il minore stretto grazie a $(*)$.

$$\begin{aligned}
\mathcal{D}_X(Y - X^+) &= \mathcal{D}_X(Y) - \mathcal{D}_X(X) + P(X) \\
&= \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) + A_1 Y \\
&- \sum_{p=2}^{n} (p-1) A_p X^p + A_0 \\
&\leq \sum_{p=2}^{n} A_p N_p < 0
\end{aligned}$$

Poiché $-D_X$ è una M-matrice si ha $Y - X^+ > 0$. □

Dimostrazione.

Dove $N_p = \sum_{q=1}^{p} X^{p-q} Y X^{q-1} - (p-1)X^p - Y^p < 0$ per il lemma (1), $A_p \geq 0$ e vale il minore stretto grazie a $(*)$.

$$\begin{aligned}
\mathcal{D}_X(Y - X^+) &= \mathcal{D}_X(Y) - \mathcal{D}_X(X) + P(X) \\
&= \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) + A_1 Y \\
&- \sum_{p=2}^{n} (p-1) A_p X^p + A_0 \\
&\leq \sum_{p=2}^{n} A_p N_p < 0
\end{aligned}$$

Poiché $-D_X$ è una M-matrice si ha $Y - X^+ > 0$.

Dimostrazione.

Dove $N_p = \sum_{q=1}^{p} X^{p-q} Y X^{q-1} - (p-1)X^p - Y^p < 0$ per il lemma (1), $A_p \geq 0$ e vale il minore stretto grazie a $(*)$.

$$\begin{aligned}\mathcal{D}_X(Y - X^+) &= \mathcal{D}_X(Y) - \mathcal{D}_X(X) + P(X) \\ &= \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) + A_1 Y \\ &\quad - \sum_{p=2}^{n} (p-1) A_p X^p + A_0 \\ &\leq \sum_{p=2}^{n} A_p N_p < 0\end{aligned}$$

Poiché $-D_X$ è una M-matrice si ha $Y - X^+ > 0$. □

Dimostrazione.

Dove $N_p = \sum_{q=1}^{p} X^{p-q} Y X^{q-1} - (p-1)X^p - Y^p < 0$ per il lemma (1), $A_p \geq 0$ e vale il minore stretto grazie a $(*)$.

$$\begin{aligned}\mathcal{D}_X(Y - X^+) &= \mathcal{D}_X(Y) - \mathcal{D}_X(X) + P(X) \\ &= \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) + A_1 Y \\ &- \sum_{p=2}^{n} (p-1) A_p X^p + A_0 \\ &\leq \sum_{p=2}^{n} A_p N_p < 0\end{aligned}$$

Poiché $-D_X$ è una M-matrice si ha $Y - X^+ > 0$.

Dimostrazione.

Dove $N_p = \sum_{q=1}^{p} X^{p-q} Y X^{q-1} - (p-1)X^p - Y^p < 0$ per il lemma (1), $A_p \geq 0$ e vale il minore stretto grazie a $(*)$.

$$\begin{aligned}
\mathcal{D}_X(Y - X^+) &= \mathcal{D}_X(Y) - \mathcal{D}_X(X) + P(X) \\
&= \sum_{p=2}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} Y X^{q-1} \Big) + A_1 Y \\
&- \sum_{p=2}^{n} (p-1) A_p X^p + A_0 \\
&\leq \sum_{p=2}^{n} A_p N_p < 0
\end{aligned}$$

Poiché $-D_X$ è una M-matrice si ha $Y - X^+ > 0$. □

Lemma (4)

Siano X e H matrici non negative e sia $\alpha \in \mathbb{R}^+$, allora

$$P(X + \alpha H) \geq P(X) + \alpha \mathcal{D}_X(H)$$

Teorema (di convergenza)

Se esiste Y una matrice positiva tale che $P(Y) \leq 0$, allora la successione $\{X_i\}$ generata dal metodo di Newton con $X_0 = 0$ e

$$\begin{cases} \mathcal{D}_{X_i}(H_i) = -P(X_i) \\ X_{i+1} = X_i + H_i \end{cases}$$

è ben definita, monotona debolmente crescente, e converge alla minima soluzione non negativa. Inoltre

$$-D_{X_i} = -\sum_{p=1}^{n} \left[\sum_{q=1}^{p} (X_i^{q-1})^T \otimes (A_p X_i^{p-q}) \right]$$

è una M-matrice non singolare ad ogni iterazione.

Teorema (di convergenza)

Se esiste Y una matrice positiva tale che $P(Y) \leq 0$, allora la successione $\{X_i\}$ generata dal metodo di Newton con $X_0 = 0$ e

$$\begin{cases} \mathcal{D}_{X_i}(H_i) = -P(X_i) \\ X_{i+1} = X_i + H_i \end{cases}$$

è ben definita, monotona debolmente crescente, e converge alla minima soluzione non negativa. Inoltre

$$-D_{X_i} = -\sum_{p=1}^{n}\left[\sum_{q=1}^{p}(X_i^{q-1})^T \otimes (A_p X_i^{p-q})\right]$$

è una M-matrice non singolare ad ogni iterazione.

Teorema (di convergenza)

Se esiste Y una matrice positiva tale che $P(Y) \leq 0$, allora la successione $\{X_i\}$ generata dal metodo di Newton con $X_0 = 0$ e

$$\begin{cases} \mathcal{D}_{X_i}(H_i) = -P(X_i) \\ X_{i+1} = X_i + H_i \end{cases}$$

è ben definita, monotona debolmente crescente, e converge alla minima soluzione non negativa. Inoltre

$$-D_{X_i} = -\sum_{p=1}^{n}\left[\sum_{q=1}^{p}(X_i^{q-1})^T \otimes (A_p X_i^{p-q})\right]$$

è una M-matrice non singolare ad ogni iterazione.

Teorema (di convergenza)

Se esiste Y una matrice positiva tale che $P(Y) \leq 0$, allora la successione $\{X_i\}$ generata dal metodo di Newton con $X_0 = 0$ e

$$\begin{cases} \mathcal{D}_{X_i}(H_i) = -P(X_i) \\ X_{i+1} = X_i + H_i \end{cases}$$

è ben definita, monotona debolmente crescente, e converge alla minima soluzione non negativa. Inoltre

$$-D_{X_i} = -\sum_{p=1}^{n}\left[\sum_{q=1}^{p}(X_i^{q-1})^T \otimes (A_p X_i^{p-q})\right]$$

è una M-matrice non singolare ad ogni iterazione.

Teorema (di convergenza)

Se esiste Y una matrice positiva tale che $P(Y) \leq 0$, allora la successione $\{X_i\}$ generata dal metodo di Newton con $X_0 = 0$ e

$$\begin{cases} \mathcal{D}_{X_i}(H_i) = -P(X_i) \\ X_{i+1} = X_i + H_i \end{cases}$$

è ben definita, monotona debolmente crescente, e converge alla minima soluzione non negativa. Inoltre

$$-D_{X_i} = -\sum_{p=1}^{n} \left[\sum_{q=1}^{p} (X_i^{q-1})^T \otimes (A_p X_i^{p-q}) \right]$$

è una M-matrice non singolare ad ogni iterazione.

Teorema (di convergenza)

Se esiste Y una matrice positiva tale che $P(Y) \leq 0$, allora la successione $\{X_i\}$ generata dal metodo di Newton con $X_0 = 0$ e

$$\begin{cases} \mathcal{D}_{X_i}(H_i) = -P(X_i) \\ X_{i+1} = X_i + H_i \end{cases}$$

è ben definita, monotona debolmente crescente, e converge alla minima soluzione non negativa. Inoltre

$$-D_{X_i} = -\sum_{p=1}^{n} \left[\sum_{q=1}^{p} (X_i^{q-1})^T \otimes (A_p X_i^{p-q}) \right]$$

è una M-matrice non singolare ad ogni iterazione.

Dimostrazione.

Dimostriamo per induzione su k i seguenti fatti:

- $X_k \leq X_{k+1}$.
- $X_k < Y$.
- $P(X_k) \geq 0$.
- $-D_{X_k}$ è una M-matrice non singolare.

Passo base $k = 0$:

- $0 = X_0 \leq (-A_1^{-1})A_0 = X_1$
- $0 = X_0 < Y$
- $P(X_0) = A_0 \geq 0$
- $-D_{X_0} = -I \otimes A_1$

Dimostrazione.

Dimostriamo per induzione su k i seguenti fatti:

- $X_k \leq X_{k+1}$.
- $X_k < Y$.
- $P(X_k) \geq 0$.
- $-D_{X_k}$ è una M-matrice non singolare.

Passo base $k = 0$:

- $0 = X_0 \leq (-A_1^{-1})A_0 = X_1$
- $0 = X_0 < Y$
- $P(X_0) = A_0 \geq 0$
- $-D_{X_0} = -I \otimes A_1$

Dimostrazione.

Dimostriamo per induzione su k i seguenti fatti:

- $X_k \leq X_{k+1}$.
- $X_k < Y$.
- $P(X_k) \geq 0$.
- $-D_{X_k}$ è una M-matrice non singolare.

Passo base $k = 0$:

- $0 = X_0 \leq (-A_1^{-1})A_0 = X_1$
- $0 = X_0 < Y$
- $P(X_0) = A_0 \geq 0$
- $-D_{X_0} = -I \otimes A_1$

Dimostrazione.

Dimostriamo per induzione su k i seguenti fatti:

- $X_k \leq X_{k+1}$.
- $X_k < Y$.
- $P(X_k) \geq 0$.
- $-D_{X_k}$ è una M-matrice non singolare.

Passo base $k = 0$:

- $0 = X_0 \leq (-A_1^{-1})A_0 = X_1$
- $0 = X_0 < Y$
- $P(X_0) = A_0 \geq 0$
- $-D_{X_0} = -I \otimes A_1$

Dimostrazione.

Dimostriamo per induzione su k i seguenti fatti:

- $X_k \leq X_{k+1}$.
- $X_k < Y$.
- $P(X_k) \geq 0$.
- $-D_{X_k}$ è una M-matrice non singolare.

Passo base $k = 0$:

- $0 = X_0 \leq (-A_1^{-1})A_0 = X_1$
- $0 = X_0 < Y$
- $P(X_0) = A_0 \geq 0$
- $-D_{X_0} = -I \otimes A_1$

Dimostrazione.

Dimostriamo per induzione su k i seguenti fatti:

- $X_k \leq X_{k+1}$.
- $X_k < Y$.
- $P(X_k) \geq 0$.
- $-D_{X_k}$ è una M-matrice non singolare.

Passo base $k = 0$:

- $0 = X_0 \leq (-A_1^{-1})A_0 = X_1$
- $0 = X_0 < Y$
- $P(X_0) = A_0 \geq 0$
- $-D_{X_0} = -I \otimes A_1$

Dimostrazione.

Passo induttivo $(k) \to (k+1)$:

- Poiché $-D_{X_k}$ è una M-matrice non singolare, per la seconda parte del lemma (3) si ha $X_{k+1} = X_k + H_k < Y$.
- Per ipotesi induttiva $X_k \leq X_{k+1}$, per cui si può applicare il lemma (4):

$$P(X_{k+1}) \geq P(X_k) + \mathcal{D}_{X_k}(H_k) = P(X_k) - P(X_k) = 0$$

- Dai due punti appena dimostrati e dalla prima parte del lemma (3) segue che $-D_{X_{k+1}}$ è una M-matrice non singolare
- $-\mathcal{D}_{X_{k+1}}(X_{k+2} - X_{k+1}) = -\mathcal{D}_{X_{k+1}}(H_{k+1}) = P(X_{k+1}) \geq 0$, poiché $-D_{X_{k+1}}$ è una M-matrice non singolare si ha $X_{k+2} - X_{k+1} \geq 0$

Dimostrazione.

Passo induttivo $(k) \to (k+1)$:

- Poiché $-D_{X_k}$ è una M-matrice non singolare, per la seconda parte del lemma (3) si ha $X_{k+1} = X_k + H_k < Y$.
- Per ipotesi induttiva $X_k \leq X_{k+1}$, per cui si può applicare il lemma (4):

$$P(X_{k+1}) \geq P(X_k) + \mathcal{D}_{X_k}(H_k) = P(X_k) - P(X_k) = 0$$

- Dai due punti appena dimostrati e dalla prima parte del lemma (3) segue che $-D_{X_{k+1}}$ è una M-matrice non singolare
- $-\mathcal{D}_{X_{k+1}}(X_{k+2} - X_{k+1}) = -\mathcal{D}_{X_{k+1}}(H_{k+1}) = P(X_{k+1}) \geq 0$, poiché $-D_{X_{k+1}}$ è una M-matrice non singolare si ha $X_{k+2} - X_{k+1} \geq 0$

Dimostrazione.

Passo induttivo $(k) \rightarrow (k+1)$:

- Poiché $-D_{X_k}$ è una M-matrice non singolare, per la seconda parte del lemma (3) si ha $X_{k+1} = X_k + H_k < Y$.
- Per ipotesi induttiva $X_k \leq X_{k+1}$, per cui si può applicare il lemma (4):

$$P(X_{k+1}) \geq P(X_k) + \mathcal{D}_{X_k}(H_k) = P(X_k) - P(X_k) = 0$$

- Dai due punti appena dimostrati e dalla prima parte del lemma (3) segue che $-D_{X_{k+1}}$ è una M-matrice non singolare
- $-\mathcal{D}_{X_{k+1}}(X_{k+2} - X_{k+1}) = -\mathcal{D}_{X_{k+1}}(H_{k+1}) = P(X_{k+1}) \geq 0$, poiché $-D_{X_{k+1}}$ è una M-matrice non singolare si ha $X_{k+2} - X_{k+1} \geq 0$

Dimostrazione.

Passo induttivo $(k) \to (k+1)$:

- Poiché $-D_{X_k}$ è una M-matrice non singolare, per la seconda parte del lemma (3) si ha $X_{k+1} = X_k + H_k < Y$.
- Per ipotesi induttiva $X_k \leq X_{k+1}$, per cui si può applicare il lemma (4):

$$P(X_{k+1}) \geq P(X_k) + \mathcal{D}_{X_k}(H_k) = P(X_k) - P(X_k) = 0$$

- Dai due punti appena dimostrati e dalla prima parte del lemma (3) segue che $-D_{X_{k+1}}$ è una M-matrice non singolare
- $-\mathcal{D}_{X_{k+1}}(X_{k+2} - X_{k+1}) = -\mathcal{D}_{X_{k+1}}(H_{k+1}) = P(X_{k+1}) \geq 0$, poiché $-D_{X_{k+1}}$ è una M-matrice non singolare si ha $X_{k+2} - X_{k+1} \geq 0$

Dimostrazione.

Passo induttivo $(k) \to (k+1)$:

- Poiché $-D_{X_k}$ è una M-matrice non singolare, per la seconda parte del lemma (3) si ha $X_{k+1} = X_k + H_k < Y$.
- Per ipotesi induttiva $X_k \leq X_{k+1}$, per cui si può applicare il lemma (4):

$$P(X_{k+1}) \geq P(X_k) + \mathcal{D}_{X_k}(H_k) = P(X_k) - P(X_k) = 0$$

- Dai due punti appena dimostrati e dalla prima parte del lemma (3) segue che $-D_{X_{k+1}}$ è una M-matrice non singolare
- $-\mathcal{D}_{X_{k+1}}(X_{k+2} - X_{k+1}) = -\mathcal{D}_{X_{k+1}}(H_{k+1}) = P(X_{k+1}) \geq 0$, poiché $-D_{X_{k+1}}$ è una M-matrice non singolare si ha $X_{k+2} - X_{k+1} \geq 0$

Dimostrazione.

Passo induttivo $(k) \to (k+1)$:

- Poiché $-D_{X_k}$ è una M-matrice non singolare, per la seconda parte del lemma (3) si ha $X_{k+1} = X_k + H_k < Y$.
- Per ipotesi induttiva $X_k \leq X_{k+1}$, per cui si può applicare il lemma (4):

$$P(X_{k+1}) \geq P(X_k) + \mathcal{D}_{X_k}(H_k) = P(X_k) - P(X_k) = 0$$

- Dai due punti appena dimostrati e dalla prima parte del lemma (3) segue che $-D_{X_{k+1}}$ è una M-matrice non singolare
- $-\mathcal{D}_{X_{k+1}}(X_{k+2} - X_{k+1}) = -\mathcal{D}_{X_{k+1}}(H_{k+1}) = P(X_{k+1}) \geq 0$, poiché $-D_{X_{k+1}}$ è una M-matrice non singolare si ha $X_{k+2} - X_{k+1} \geq 0$

Dimostrazione.

Passo induttivo $(k) \to (k+1)$:

- Poiché $-D_{X_k}$ è una M-matrice non singolare, per la seconda parte del lemma (3) si ha $X_{k+1} = X_k + H_k < Y$.
- Per ipotesi induttiva $X_k \leq X_{k+1}$, per cui si può applicare il lemma (4):

$$P(X_{k+1}) \geq P(X_k) + \mathcal{D}_{X_k}(H_k) = P(X_k) - P(X_k) = 0$$

- Dai due punti appena dimostrati e dalla prima parte del lemma (3) segue che $-D_{X_{k+1}}$ è una M-matrice non singolare
- $-\mathcal{D}_{X_{k+1}}(X_{k+2} - X_{k+1}) = -\mathcal{D}_{X_{k+1}}(H_{k+1}) = P(X_{k+1}) \geq 0$, poiché $-D_{X_{k+1}}$ è una M-matrice non singolare si ha $X_{k+2} - X_{k+1} \geq 0$

Dimostrazione.

Passo induttivo $(k) \to (k+1)$:

- Poiché $-D_{X_k}$ è una M-matrice non singolare, per la seconda parte del lemma (3) si ha $X_{k+1} = X_k + H_k < Y$.
- Per ipotesi induttiva $X_k \leq X_{k+1}$, per cui si può applicare il lemma (4):

$$P(X_{k+1}) \geq P(X_k) + \mathcal{D}_{X_k}(H_k) = P(X_k) - P(X_k) = 0$$

- Dai due punti appena dimostrati e dalla prima parte del lemma (3) segue che $-D_{X_{k+1}}$ è una M-matrice non singolare
- $-\mathcal{D}_{X_{k+1}}(X_{k+2} - X_{k+1}) = -\mathcal{D}_{X_{k+1}}(H_{k+1}) = P(X_{k+1}) \geq 0$, poiché $-D_{X_{k+1}}$ è una M-matrice non singolare si ha $X_{k+2} - X_{k+1} \geq 0$

Dimostrazione.

Passo induttivo $(k) \to (k+1)$:

- Poiché $-D_{X_k}$ è una M-matrice non singolare, per la seconda parte del lemma (3) si ha $X_{k+1} = X_k + H_k < Y$.
- Per ipotesi induttiva $X_k \leq X_{k+1}$, per cui si può applicare il lemma (4):

$$P(X_{k+1}) \geq P(X_k) + \mathcal{D}_{X_k}(H_k) = P(X_k) - P(X_k) = 0$$

- Dai due punti appena dimostrati e dalla prima parte del lemma (3) segue che $-D_{X_{k+1}}$ è una M-matrice non singolare
- $-\mathcal{D}_{X_{k+1}}(X_{k+2} - X_{k+1}) = -\mathcal{D}_{X_{k+1}}(H_{k+1}) = P(X_{k+1}) \geq 0$, poiché $-D_{X_{k+1}}$ è una M-matrice non singolare si ha $X_{k+2} - X_{k+1} \geq 0$

Dimostrazione.

La successione $\{X_k\}$ è crescente e limitata, quindi esiste una matrice L tale che $L := \lim_{k\to\infty} X_k$.

Osserviamo che $H_k = X_{k+1} - X_k \to L - L = 0$, consideriamo la relazione $\mathcal{D}_{X_k}(H_k) = -P(X_k)$ e andiamo al limite a destra e sinistra:

- $-P(X_k) \to -P(L)$
- $\mathcal{D}_{X_k}(H_k) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X_k^{p-q} H_k X_k^{q-1} \Big) \to 0$

Quindi $P(L) = 0$. Sia ora L' un'altra soluzione, applicando il corollario (2) con $Y = L'$ e $X = X_k$, si ottiene per induzione che $X_k \leq L'$ per ogni k e si conclude che $L \leq L'$. $\square$

Dimostrazione.

La successione $\{X_k\}$ è crescente e limitata, quindi esiste una matrice L tale che $L := \lim_{k\to\infty} X_k$.
Osserviamo che $H_k = X_{k+1} - X_k \to L - L = 0$, consideriamo la relazione $\mathcal{D}_{X_k}(H_k) = -P(X_k)$ e andiamo al limite a destra e sinistra:

- $-P(X_k) \to -P(L)$
- $\mathcal{D}_{X_k}(H_k) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X_k^{p-q} H_k X_k^{q-1} \Big) \to 0$

Quindi $P(L) = 0$. Sia ora L' un'altra soluzione, applicando il corollario (2) con $Y = L'$ e $X = X_k$, si ottiene per induzione che $X_k \le L'$ per ogni k e si conclude che $L \le L'$. □

Dimostrazione.

La successione $\{X_k\}$ è crescente e limitata, quindi esiste una matrice L tale che $L := \lim_{k\to\infty} X_k$.
Osserviamo che $H_k = X_{k+1} - X_k \to L - L = 0$, consideriamo la relazione $\mathcal{D}_{X_k}(H_k) = -P(X_k)$ e andiamo al limite a destra e sinistra:

- $-P(X_k) \to -P(L)$
- $\mathcal{D}_{X_k}(H_k) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X_k^{p-q} H_k X_k^{q-1} \Big) \to 0$

Quindi $P(L) = 0$. Sia ora L' un'altra soluzione, applicando il corollario (2) con $Y = L'$ e $X = X_k$, si ottiene per induzione che $X_k \le L'$ per ogni k e si conclude che $L \le L'$. □

Dimostrazione.

La successione $\{X_k\}$ è crescente e limitata, quindi esiste una matrice L tale che $L := \lim_{k\to\infty} X_k$.
Osserviamo che $H_k = X_{k+1} - X_k \to L - L = 0$, consideriamo la relazione $\mathcal{D}_{X_k}(H_k) = -P(X_k)$ e andiamo al limite a destra e sinistra:

- $-P(X_k) \to -P(L)$
- $\mathcal{D}_{X_k}(H_k) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X_k^{p-q} H_k X_k^{q-1} \Big) \to 0$

Quindi $P(L) = 0$. Sia ora L' un'altra soluzione, applicando il corollario (2) con $Y = L'$ e $X = X_k$, si ottiene per induzione che $X_k \leq L'$ per ogni k e si conclude che $L \leq L'$. $\square$

Dimostrazione.

La successione $\{X_k\}$ è crescente e limitata, quindi esiste una matrice L tale che $L := \lim_{k\to\infty} X_k$.
Osserviamo che $H_k = X_{k+1} - X_k \to L - L = 0$, consideriamo la relazione $\mathcal{D}_{X_k}(H_k) = -P(X_k)$ e andiamo al limite a destra e sinistra:

- $-P(X_k) \to -P(L)$
- $\mathcal{D}_{X_k}(H_k) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X_k^{p-q} H_k X_k^{q-1} \Big) \to 0$

Quindi $P(L) = 0$. Sia ora L' un'altra soluzione, applicando il corollario (2) con $Y = L'$ e $X = X_k$, si ottiene per induzione che $X_k \le L'$ per ogni k e si conclude che $L \le L'$. $\square$

Dimostrazione.

La successione $\{X_k\}$ è crescente e limitata, quindi esiste una matrice L tale che $L := \lim_{k\to\infty} X_k$.
Osserviamo che $H_k = X_{k+1} - X_k \to L - L = 0$, consideriamo la relazione $\mathcal{D}_{X_k}(H_k) = -P(X_k)$ e andiamo al limite a destra e sinistra:

- $-P(X_k) \to -P(L)$
- $\mathcal{D}_{X_k}(H_k) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X_k^{p-q} H_k X_k^{q-1} \Big) \to 0$

Quindi $P(L) = 0$. Sia ora L' un'altra soluzione, applicando il corollario (2) con $Y = L'$ e $X = X_k$, si ottiene per induzione che $X_k \leq L'$ per ogni k e si conclude che $L \leq L'$. □

Dimostrazione.

La successione $\{X_k\}$ è crescente e limitata, quindi esiste una matrice L tale che $L := \lim_{k\to\infty} X_k$.
Osserviamo che $H_k = X_{k+1} - X_k \to L - L = 0$, consideriamo la relazione $\mathcal{D}_{X_k}(H_k) = -P(X_k)$ e andiamo al limite a destra e sinistra:

- $-P(X_k) \to -P(L)$
- $\mathcal{D}_{X_k}(H_k) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X_k^{p-q} H_k X_k^{q-1} \Big) \to 0$

Quindi $P(L) = 0$. Sia ora L' un'altra soluzione, applicando il corollario (2) con $Y = L'$ e $X = X_k$, si ottiene per induzione che $X_k \leq L'$ per ogni k e si conclude che $L \leq L'$. $\square$

Dimostrazione.

La successione $\{X_k\}$ è crescente e limitata, quindi esiste una matrice L tale che $L := \lim_{k\to\infty} X_k$.
Osserviamo che $H_k = X_{k+1} - X_k \to L - L = 0$, consideriamo la relazione $\mathcal{D}_{X_k}(H_k) = -P(X_k)$ e andiamo al limite a destra e sinistra:

- $-P(X_k) \to -P(L)$
- $\mathcal{D}_{X_k}(H_k) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X_k^{p-q} H_k X_k^{q-1} \Big) \to 0$

Quindi $P(L) = 0$. Sia ora L' un'altra soluzione, applicando il corollario (2) con $Y = L'$ e $X = X_k$, si ottiene per induzione che $X_k \leq L'$ per ogni k e si conclude che $L \leq L'$.

Dimostrazione.

La successione $\{X_k\}$ è crescente e limitata, quindi esiste una matrice L tale che $L := \lim_{k\to\infty} X_k$.

Osserviamo che $H_k = X_{k+1} - X_k \to L - L = 0$, consideriamo la relazione $\mathcal{D}_{X_k}(H_k) = -P(X_k)$ e andiamo al limite a destra e sinistra:

- $-P(X_k) \to -P(L)$
- $\mathcal{D}_{X_k}(H_k) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X_k^{p-q} H_k X_k^{q-1} \Big) \to 0$

Quindi $P(L) = 0$. Sia ora L' un'altra soluzione, applicando il corollario (2) con $Y = L'$ e $X = X_k$, si ottiene per induzione che $X_k \leq L'$ per ogni k e si conclude che $L \leq L'$.

Dimostrazione.

La successione $\{X_k\}$ è crescente e limitata, quindi esiste una matrice L tale che $L := \lim_{k\to\infty} X_k$.
Osserviamo che $H_k = X_{k+1} - X_k \to L - L = 0$, consideriamo la relazione $\mathcal{D}_{X_k}(H_k) = -P(X_k)$ e andiamo al limite a destra e sinistra:

- $-P(X_k) \to -P(L)$
- $\mathcal{D}_{X_k}(H_k) = \sum_{p=1}^{n} A_p \left(\sum_{q=1}^{p} X_k^{p-q} H_k X_k^{q-1} \right) \to 0$

Quindi $P(L) = 0$. Sia ora L' un'altra soluzione, applicando il corollario (2) con $Y = L'$ e $X = X_k$, si ottiene per induzione che $X_k \leq L'$ per ogni k e si conclude che $L \leq L'$. $\square$

Dimostrazione.

La successione $\{X_k\}$ è crescente e limitata, quindi esiste una matrice L tale che $L := \lim_{k\to\infty} X_k$.
Osserviamo che $H_k = X_{k+1} - X_k \to L - L = 0$, consideriamo la relazione $\mathcal{D}_{X_k}(H_k) = -P(X_k)$ e andiamo al limite a destra e sinistra:

- $-P(X_k) \to -P(L)$
- $\mathcal{D}_{X_k}(H_k) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X_k^{p-q} H_k X_k^{q-1} \Big) \to 0$

Quindi $P(L) = 0$. Sia ora L' un'altra soluzione, applicando il corollario (2) con $Y = L'$ e $X = X_k$, si ottiene per induzione che $X_k \leq L'$ per ogni k e si conclude che $L \leq L'$.

Dimostrazione.

La successione $\{X_k\}$ è crescente e limitata, quindi esiste una matrice L tale che $L := \lim_{k\to\infty} X_k$.
Osserviamo che $H_k = X_{k+1} - X_k \to L - L = 0$, consideriamo la relazione $\mathcal{D}_{X_k}(H_k) = -P(X_k)$ e andiamo al limite a destra e sinistra:

- $-P(X_k) \to -P(L)$
- $\mathcal{D}_{X_k}(H_k) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X_k^{p-q} H_k X_k^{q-1} \Big) \to 0$

Quindi $P(L) = 0$. Sia ora L' un'altra soluzione, applicando il corollario (2) con $Y = L'$ e $X = X_k$, si ottiene per induzione che $X_k \leq L'$ per ogni k e si conclude che $L \leq L'$. $\square$

Corollario

Vale allo stesso modo la tesi del teorema di convergenza se X_0 *è una qualsiasi matrice non negativa tale che* $X_0 < Y$ *e* $P(X_0) \geq 0$.

Metodo di Newton rilassato

Ispirandoci ai metodi di Newton con ricerca esatta usati nei problemi di ottimizzazione, cerchiamo un multiplo del passo di Newton che minimizzi la funzione di merito:

$$p(\lambda) = \|P(X_i + \lambda H_i)\|^2.$$

Osservazioni (Punti deboli della ricerca esatta)

- Non è garantito che $-D_{X_i}$ sia una M-matrice non singolare.
- Non è computazionalmente semplice calcolare la funzione di merito e il suo minimo quando il grado di $P(X)$ è alto.

Il seguente teorema mostra una formula semplice per trovare una direzione λ che velocizzi la convergenza del metodo di Newton e che ne conservi le proprietà.

Metodo di Newton rilassato

Ispirandoci ai metodi di Newton con ricerca esatta usati nei problemi di ottimizzazione, cerchiamo un multiplo del passo di Newton che minimizzi la funzione di merito:

$$p(\lambda) = \|P(X_i + \lambda H_i)\|^2.$$

Osservazioni (Punti deboli della ricerca esatta)

- Non è garantito che $-D_{X_i}$ sia una M-matrice non singolare.
- Non è computazionalmente semplice calcolare la funzione di merito e il suo minimo quando il grado di $P(X)$ è alto.

Il seguente teorema mostra una formula semplice per trovare una direzione λ che velocizzi la convergenza del metodo di Newton e che ne conservi le proprietà.

Metodo di Newton rilassato

Ispirandoci ai metodi di Newton con ricerca esatta usati nei problemi di ottimizzazione, cerchiamo un multiplo del passo di Newton che minimizzi la funzione di merito:

$$p(\lambda) = \|P(X_i + \lambda H_i)\|^2.$$

Osservazioni (Punti deboli della ricerca esatta)

- Non è garantito che $-D_{X_i}$ sia una M-matrice non singolare.
- Non è computazionalmente semplice calcolare la funzione di merito e il suo minimo quando il grado di $P(X)$ è alto.

Il seguente teorema mostra una formula semplice per trovare una direzione λ che velocizzi la convergenza del metodo di Newton e che ne conservi le proprietà.

Metodo di Newton rilassato

Ispirandoci ai metodi di Newton con ricerca esatta usati nei problemi di ottimizzazione, cerchiamo un multiplo del passo di Newton che minimizzi la funzione di merito:

$$p(\lambda) = \|P(X_i + \lambda H_i)\|^2.$$

Osservazioni (Punti deboli della ricerca esatta)

- Non è garantito che $-D_{X_i}$ sia una M-matrice non singolare.
- Non è computazionalmente semplice calcolare la funzione di merito e il suo minimo quando il grado di $P(X)$ è alto.

Il seguente teorema mostra una formula semplice per trovare una direzione λ che velocizzi la convergenza del metodo di Newton e che ne conservi le proprietà.

Metodo di Newton rilassato

Ispirandoci ai metodi di Newton con ricerca esatta usati nei problemi di ottimizzazione, cerchiamo un multiplo del passo di Newton che minimizzi la funzione di merito:

$$p(\lambda) = \|P(X_i + \lambda H_i)\|^2.$$

Osservazioni (Punti deboli della ricerca esatta)

- Non è garantito che $-D_{X_i}$ sia una M-matrice non singolare.
- Non è computazionalmente semplice calcolare la funzione di merito e il suo minimo quando il grado di $P(X)$ è alto.

Il seguente teorema mostra una formula semplice per trovare una direzione λ che velocizzi la convergenza del metodo di Newton e che ne conservi le proprietà.

Metodo di Newton rilassato

Ispirandoci ai metodi di Newton con ricerca esatta usati nei problemi di ottimizzazione, cerchiamo un multiplo del passo di Newton che minimizzi la funzione di merito:

$$p(\lambda) = \|P(X_i + \lambda H_i)\|^2.$$

Osservazioni (Punti deboli della ricerca esatta)

- Non è garantito che $-D_{X_i}$ sia una M-matrice non singolare.
- Non è computazionalmente semplice calcolare la funzione di merito e il suo minimo quando il grado di $P(X)$ è alto.

Il seguente teorema mostra una formula semplice per trovare una direzione λ che velocizzi la convergenza del metodo di Newton e che ne conservi le proprietà.

Proposizione

Se esiste una Y matrice non negativa tale che $P(Y) \leq 0$, per ogni X tale che $Y > X \geq 0$ e $P(X) \geq 0$ definiamo λ e H tramite $\mathcal{D}_X(H) = -P(X)$ e

$$\lambda := \min \left\{ \frac{[P(X)]_{ij} + [P(X^+)]_{ij}}{[P(X)]_{ij}} : [P(X)]_{ij} > 0 \right\}.$$

Allora valgono i seguenti fatti:

- $X^+ \leq X + \lambda H < Y$.
- $P(X + \lambda H) \geq 0$.

Proposizione

Se esiste una Y matrice non negativa tale che $P(Y) \leq 0$, per ogni X tale che $Y > X \geq 0$ e $P(X) \geq 0$ definiamo λ e H tramite $\mathcal{D}_X(H) = -P(X)$ e

$$\lambda := \min \left\{ \frac{[P(X)]_{ij} + [P(X^+)]_{ij}}{[P(X)]_{ij}} : [P(X)]_{ij} > 0 \right\}.$$

Allora valgono i seguenti fatti:

- $X^+ \leq X + \lambda H < Y$.
- $P(X + \lambda H) \geq 0$.

Proposizione

Se esiste una Y matrice non negativa tale che $P(Y) \leq 0$, per ogni X tale che $Y > X \geq 0$ e $P(X) \geq 0$ definiamo λ e H tramite $\mathcal{D}_X(H) = -P(X)$ *e*

$$\lambda := \min \left\{ \frac{[P(X)]_{ij} + [P(X^+)]_{ij}}{[P(X)]_{ij}} : [P(X)]_{ij} > 0 \right\}.$$

Allora valgono i seguenti fatti:

- $X^+ \leq X + \lambda H < Y$.
- $P(X + \lambda H) \geq 0$.

Proposizione

Se esiste una Y matrice non negativa tale che $P(Y) \leq 0$, per ogni X tale che $Y > X \geq 0$ e $P(X) \geq 0$ definiamo λ e H tramite $\mathcal{D}_X(H) = -P(X)$ e

$$\lambda := \min \left\{ \frac{[P(X)]_{ij} + [P(X^+)]_{ij}}{[P(X)]_{ij}} : [P(X)]_{ij} > 0 \right\}.$$

Allora valgono i seguenti fatti:

- $X^+ \leq X + \lambda H < Y$.
- $P(X + \lambda H) \geq 0$.

Proposizione

Se esiste una Y *matrice non negativa tale che* $P(Y) \leq 0$*, per ogni* X *tale che* $Y > X \geq 0$ *e* $P(X) \geq 0$ *definiamo* λ *e* H *tramite* $\mathcal{D}_X(H) = -P(X)$ *e*

$$\lambda := \min \left\{ \frac{[P(X)]_{ij} + [P(X^+)]_{ij}}{[P(X)]_{ij}} : [P(X)]_{ij} > 0 \right\}.$$

Allora valgono i seguenti fatti:

- $X^+ \leq X + \lambda H < Y$.
- $P(X + \lambda H) \geq 0$.

Proposizione

Se esiste una Y matrice non negativa tale che $P(Y) \leq 0$, per ogni X tale che $Y > X \geq 0$ e $P(X) \geq 0$ definiamo λ e H tramite $\mathcal{D}_X(H) = -P(X)$ e

$$\lambda := \min \left\{ \frac{[P(X)]_{ij} + [P(X^+)]_{ij}}{[P(X)]_{ij}} : [P(X)]_{ij} > 0 \right\}.$$

Allora valgono i seguenti fatti:

- $X^+ \leq X + \lambda H < Y$.
- $P(X + \lambda H) \geq 0$.

Proposizione

Se esiste una Y *matrice non negativa tale che* $P(Y) \leq 0$*, per ogni* X *tale che* $Y > X \geq 0$ *e* $P(X) \geq 0$ *definiamo* λ *e* H *tramite* $\mathcal{D}_X(H) = -P(X)$ *e*

$$\lambda := \min \left\{ \frac{[P(X)]_{ij} + [P(X^+)]_{ij}}{[P(X)]_{ij}} : [P(X)]_{ij} > 0 \right\}.$$

Allora valgono i seguenti fatti:

- $X^+ \leq X + \lambda H < Y$.
- $P(X + \lambda H) \geq 0$.

Osservazioni

- Poiché $X_\lambda \geq X^+$, dal lemma (2) si ha che $\mathcal{D}_{X_\lambda}(H) \geq \mathcal{D}_{X^+}(H)$, quindi $P(X^+) \geq P(X_\lambda)$ per cui il metodo di Newton rilassato può essere visto come un metodo discendente per minimizzare $\|P(X)\|$.
- Dalla proposizione precedente e dal lemma (3) segue che $-\mathcal{D}_{X_\lambda}$ è una M-matrice.

Osservazioni

- Poiché $X_\lambda \geq X^+$, dal lemma (2) si ha che $\mathcal{D}_{X_\lambda}(H) \geq \mathcal{D}_{X^+}(H)$, quindi $P(X^+) \geq P(X_\lambda)$ per cui il metodo di Newton rilassato può essere visto come un metodo discendente per minimizzare $\|P(X)\|$.
- Dalla proposizione precedente e dal lemma (3) segue che $-\mathcal{D}_{X_\lambda}$ è una M-matrice.

Osservazioni

- Poiché $X_\lambda \geq X^+$, dal lemma (2) si ha che $\mathcal{D}_{X_\lambda}(H) \geq \mathcal{D}_{X^+}(H)$, quindi $P(X^+) \geq P(X_\lambda)$ per cui il metodo di Newton rilassato può essere visto come un metodo discendente per minimizzare $\|P(X)\|$.
- Dalla proposizione precedente e dal lemma (3) segue che $-\mathcal{D}_{X_\lambda}$ è una M-matrice.

Teorema (di convergenza)

Se esiste Y una matrice positiva tale che $P(Y) \leq 0$, allora la successione $\{X_i\}$ generata dal metodo di Newton rilassato con $X_0 = 0$ e

$$\begin{cases} \mathcal{D}_{X_i}(H_i) = -P(X_i) \\ X_{i+1} = X_i + \lambda_i H_i \end{cases}$$

con λ_i definito come nel teorema precedente, è ben definita, monotona debolmente crescente, e converge alla minima soluzione non negativa. Inoltre $-D_{X_i+\lambda_i H_i}$ è una M-matrice non singolare ad ogni iterazione.

Metodo di Schur

Un aspetto su cui il nostro algoritmo può essere migliorato è la soluzione di $\mathcal{D}_X(H) = -P(X)$.

A questo scopo consideriamo $X = QR\overline{Q}^T$, la decomposizione di Schur di X, vale che

$$\mathcal{D}_X(H) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} H X^{q-1} \Big) = \sum_{p=1}^{n} \Big(\sum_{q=1}^{p} A_q X^{q-1} H X^{p-1} \Big) = \sum_{p=1}^{n} B_p H' R^{p-1} \overline{Q}^T$$

dove $H' = HQ$ e $B_p = \sum_{q=1}^{p} A_q X^{q-1}$.

Metodo di Schur

Un aspetto su cui il nostro algoritmo può essere migliorato è la soluzione di $\mathcal{D}_X(H) = -P(X)$.

A questo scopo consideriamo $X = QR\overline{Q}^T$, la decomposizione di Schur di X, vale che

$$\mathcal{D}_X(H) = \sum_{p=1}^{n} A_p\Big(\sum_{q=1}^{p} X^{p-q}HX^{q-1}\Big)$$
$$= \sum_{p=1}^{n}\Big(\sum_{q=1}^{p} A_q X^{q-1} H X^{p-1}\Big)$$
$$= \sum_{p=1}^{n} B_p H' R^{p-1}\overline{Q}^T$$

dove $H' = HQ$ e $B_p = \sum_{q=1}^{p} A_q X^{q-1}$.

Metodo di Schur

Un aspetto su cui il nostro algoritmo può essere migliorato è la soluzione di $\mathcal{D}_X(H) = -P(X)$.

A questo scopo consideriamo $X = QR\overline{Q}^T$, la decomposizione di Schur di X, vale che

$$\begin{aligned}\mathcal{D}_X(H) &= \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} H X^{q-1} \Big) \\ &= \sum_{p=1}^{n} \Big(\sum_{q=1}^{p} A_q X^{q-1} H X^{p-1} \Big) \\ &= \sum_{p=1}^{n} B_p H' R^{p-1} \overline{Q}^T\end{aligned}$$

dove $H' = HQ$ e $B_p = \sum_{q=1}^{p} A_q X^{q-1}$.

Metodo di Schur

Un aspetto su cui il nostro algoritmo può essere migliorato è la soluzione di $\mathcal{D}_X(H) = -P(X)$.

A questo scopo consideriamo $X = QR\overline{Q}^T$, la decomposizione di Schur di X, vale che

$$\mathcal{D}_X(H) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} H X^{q-1} \Big)$$

$$= \sum_{p=1}^{n} \Big(\sum_{q=1}^{p} A_q X^{q-1} H X^{p-1} \Big)$$

$$= \sum_{p=1}^{n} B_p H' R^{p-1} \overline{Q}^T$$

dove $H' = HQ$ e $B_p = \sum_{q=1}^{p} A_q X^{q-1}$.

Metodo di Schur

Un aspetto su cui il nostro algoritmo può essere migliorato è la soluzione di $\mathcal{D}_X(H) = -P(X)$.
A questo scopo consideriamo $X = QR\overline{Q}^T$, la decomposizione di Schur di X, vale che

$$\mathcal{D}_X(H) = \sum_{p=1}^{n} A_p \Big(\sum_{q=1}^{p} X^{p-q} H X^{q-1} \Big)$$
$$= \sum_{p=1}^{n} \Big(\sum_{q=1}^{p} A_q X^{q-1} H X^{p-1} \Big)$$
$$= \sum_{p=1}^{n} B_p H' R^{p-1} \overline{Q}^T$$

dove $H' = HQ$ e $B_p = \sum_{q=1}^{p} A_q X^{q-1}$.

Metodo di Schur

Un aspetto su cui il nostro algoritmo può essere migliorato è la soluzione di $\mathcal{D}_X(H) = -P(X)$.

A questo scopo consideriamo $X = QR\overline{Q}^T$, la decomposizione di Schur di X, vale che

$$\begin{aligned}\mathcal{D}_X(H) &= \sum_{p=1}^{n} A_p\Big(\sum_{q=1}^{p} X^{p-q}HX^{q-1}\Big)\\ &= \sum_{p=1}^{n}\Big(\sum_{q=1}^{p} A_q X^{q-1}HX^{p-1}\Big)\\ &= \sum_{p=1}^{n} B_p H' R^{p-1}\overline{Q}^T\end{aligned}$$

dove $H' = HQ$ e $B_p = \sum_{q=1}^{p} A_q X^{q-1}$.

Metodo di Schur

Un aspetto su cui il nostro algoritmo può essere migliorato è la soluzione di $\mathcal{D}_X(H) = -P(X)$.
A questo scopo consideriamo $X = QR\overline{Q}^T$, la decomposizione di Schur di X, vale che

$$\begin{aligned}\mathcal{D}_X(H) &= \sum_{p=1}^{n} A_p\Big(\sum_{q=1}^{p} X^{p-q}HX^{q-1}\Big)\\ &= \sum_{p=1}^{n}\Big(\sum_{q=1}^{p} A_q X^{q-1}HX^{p-1}\Big)\\ &= \sum_{p=1}^{n} B_p H' R^{p-1}\overline{Q}^T\end{aligned}$$

dove $H' = HQ$ e $B_p = \sum_{q=1}^{p} A_q X^{q-1}$.

Per quanto visto possiamo scrivere la nostra equazione come

$$\sum_{p=1}^{n} B_p H' R^{p-1} = -P(X)Q =: C$$

che, tramite l'operatore *vec* e il prodotto di Kronecker, diventa:

$$Mvec(H') = vec(C)$$

con $M = \sum_{p=1}^{n} (R^T)^{p-1} \otimes B_p$. R è triangolare superiore, quindi:

$$M = \begin{bmatrix} M_{11} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ M_{m1} & \dots & \dots & M_{mm} \end{bmatrix}$$

Per quanto visto possiamo scrivere la nostra equazione come

$$\sum_{p=1}^{n} B_p H' R^{p-1} = -P(X)Q =: C$$

che, tramite l'operatore *vec* e il prodotto di Kronecker, diventa:

$$Mvec(H') = vec(C)$$

con $M = \sum_{p=1}^{n} (R^T)^{p-1} \otimes B_p$. R è triangolare superiore, quindi:

$$M = \begin{bmatrix} M_{11} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ M_{m1} & \dots & \dots & M_{mm} \end{bmatrix}$$

Per quanto visto possiamo scrivere la nostra equazione come

$$\sum_{p=1}^{n} B_p H' R^{p-1} = -P(X)Q =: C$$

che, tramite l'operatore *vec* e il prodotto di Kronecker, diventa:

$$Mvec(H') = vec(C)$$

con $M = \sum_{p=1}^{n} (R^T)^{p-1} \otimes B_p$. R è triangolare superiore, quindi:

$$M = \begin{bmatrix} M_{11} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ M_{m1} & \dots & \dots & M_{mm} \end{bmatrix}$$

Per quanto visto possiamo scrivere la nostra equazione come

$$\sum_{p=1}^{n} B_p H' R^{p-1} = -P(X)Q =: C$$

che, tramite l'operatore *vec* e il prodotto di Kronecker, diventa:

$$Mvec(H') = vec(C)$$

con $M = \sum_{p=1}^{n} (R^T)^{p-1} \otimes B_p$. R è triangolare superiore, quindi:

$$M = \begin{bmatrix} M_{11} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ M_{m1} & \dots & \dots & M_{mm} \end{bmatrix}$$

Ogni blocco di M può essere scritto come

$$M_{ij} = \sum_{p=1}^{n} [R^{p-1}]_{ij} B_p$$

e si possono calcolare le colonne h'_i di H' risolvendo i seguenti sistemi lineari:

$$\begin{cases} M_{11} h'_1 = c_1 \\ M_{22} h'_2 = c_2 - M_{21} h'_1 \\ \vdots \\ M_{mm} h'_n = c_m - M_{m1} h'_1 - \cdots - M_{mm-1} h'_{m-1} \end{cases}$$

dove c_i rappresentano le colonne della matrice C.
Pertanto ci siamo ridotti a dover risolvere m sistemi di taglia m invece di uno solo di taglia m^2.

Ogni blocco di M può essere scritto come

$$M_{ij} = \sum_{p=1}^{n} [R^{p-1}]_{ij} B_p$$

e si possono calcolare le colonne h'_i di H' risolvendo i seguenti sistemi lineari:

$$\begin{cases} M_{11} h'_1 = c_1 \\ M_{22} h'_2 = c_2 - M_{21} h'_1 \\ \vdots \\ M_{mm} h'_n = c_m - M_{m1} h'_1 - \cdots - M_{mm-1} h'_{m-1} \end{cases}$$

dove c_i rappresentano le colonne della matrice C.
Pertanto ci siamo ridotti a dover risolvere m sistemi di taglia m invece di uno solo di taglia m^2.

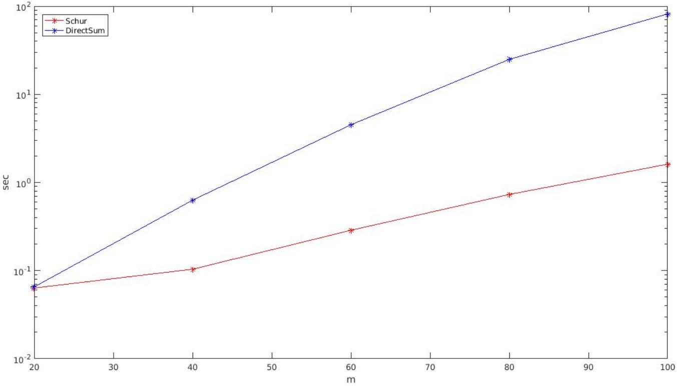
Ogni blocco di M può essere scritto come

$$M_{ij} = \sum_{p=1}^{n} [R^{p-1}]_{ij} B_p$$

e si possono calcolare le colonne h_i' di H' risolvendo i seguenti sistemi lineari:

$$\begin{cases} M_{11} h_1' = c_1 \\ M_{22} h_2' = c_2 - M_{21} h_1' \\ \vdots \\ M_{mm} h_n' = c_m - M_{m1} h_1' - \cdots - M_{mm-1} h_{m-1}' \end{cases}$$

dove c_i rappresentano le colonne della matrice C.
Pertanto ci siamo ridotti a dover risolvere m sistemi di taglia m invece di uno solo di taglia m^2.


Schur
DirectSum
sec
m
10^2
10^1
10^0
10^{-1}
10^{-2}
20
30
40
50
60
70
80
90
100