Nonlinear Matrix Equations

Based on collaboration with

Salah M. El-Sayed

SIAM Journal Matrix Analysis and Applications

Martine Reurings (PhD student VU)

Linear Algebra and its Applications

Electronic Linear Algebra

Matrix equations from systems and control theory

Discrete algebraic Riccati equation (DARE)

$$X = A^*XA + Q - A^*XB(R + B^*XB)^{-1}B^*XA$$

Matrix equations from systems and control theory

Discrete algebraic Riccati equation (DARE)

$$X = A^*XA + Q - A^*XB(R + B^*XB)^{-1}B^*XA$$

Generalization

$$X = A^*XA + Q - A^*XB(R + B^*XB)^{-1}B^*XA + \Pi(X)$$

Matrix equations from systems and control theory

Discrete algebraic Riccati equation (DARE)

$$X = A^*XA + Q - A^*XB(R + B^*XB)^{-1}B^*XA$$

Generalization

$$X = A^*XA + Q - A^*XB(R + B^*XB)^{-1}B^*XA + \Pi(X)$$

Special case

$$X + A^*X^{-1}A = Q$$

Matrix equations from systems and control theory

Discrete algebraic Riccati equation (DARE)

$$X = A^*XA + Q - A^*XB(R + B^*XB)^{-1}B^*XA$$

Generalization

$$X = A^*XA + Q - A^*XB(R + B^*XB)^{-1}B^*XA + \Pi(X)$$

Special case

$$X + A^*X^{-1}A = Q$$

Stein equation

$$X - A^*XA = Q.$$

Motivates looking at

$$X + A^*\mathcal{F}(X)A = Q$$

where

Motivates looking at

$$X + A^*\mathcal{F}(X)A = Q$$

where

- A is an $n \times n$ matrix

Motivates looking at

$$X + A^*\mathcal{F}(X)A = Q$$

where

- A is an $n \times n$ matrix

- Q positive definite and

Motivates looking at

$$X + A^*\mathcal{F}(X)A = Q$$

where

- A is an $n \times n$ matrix
- Q positive definite and
- $\mathcal{F}$ maps the positive definite matrices into the positive definite matrices and is order preserving

  $X \leq Y$ implies $\mathcal{F}(X) \leq \mathcal{F}(Y)$.

The order preserving case

$\mathcal{F}$ order preserving.

With $\mathcal{F}$ construct the map $\mathcal{G}$

$$\mathcal{G}(X) = Q - A^*\mathcal{F}(X)A.$$

Note that a solution of $X + A^*\mathcal{F}(X)A = Q$ is a fixed point of $\mathcal{G}$.

The order preserving case

$\mathcal{F}$ order preserving.

With $\mathcal{F}$ construct the map $\mathcal{G}$

$$\mathcal{G}(X) = Q - A^*\mathcal{F}(X)A.$$

Note that a solution of $X + A^*\mathcal{F}(X)A = Q$ is a fixed point of $\mathcal{G}$.

We assume throughout that

$$\mathcal{G}(Q) > 0.$$

The order preserving case

$\mathcal{F}$ order preserving.

With $\mathcal{F}$ construct the map $\mathcal{G}$

$$\mathcal{G}(X) = Q - A^*\mathcal{F}(X)A.$$

Note that a solution of $X + A^*\mathcal{F}(X)A = Q$ is a fixed point of $\mathcal{G}$.

We assume throughout that

$$\mathcal{G}(Q) > 0.$$

$\mathcal{F}$ order preserving implies $\mathcal{G}$ order reversing.

Consider

$$[\mathcal{G}(Q), Q] = \{X \mid \mathcal{G}(Q) \le X \le Q\}$$

As $\mathcal{F}(X) \ge 0$ for $X > 0$ we have $\mathcal{G}(X) \le Q$ for $X > 0$.

So this compact, convex set is non-empty, and is mapped into itself by $\mathcal{G}$.

Consider

$$[\mathcal{G}(Q), Q] = \{X \mid \mathcal{G}(Q) \leq X \leq Q\}$$

As $\mathcal{F}(X) \geq 0$ for $X > 0$ we have $\mathcal{G}(X) \leq Q$ for $X > 0$.

So this compact, convex set is non-empty, and is mapped into itself by $\mathcal{G}$.

If we assume that $\mathcal{F}$ and hence $\mathcal{G}$ is continuous then we may conclude that there is a solution from Schauder's fixed point theorem.

Non-constructive!

Consider the sequence of matrices

$$\mathcal{G}^k(Q).$$

From $\mathcal{G}(Q) \leq Q$ we have

$$\mathcal{G}(Q) \leq \mathcal{G}^2(Q)$$

as $\mathcal{G}$ is order reversing.

Also $\mathcal{G}^2$ is order preserving.

Apply $\mathcal{G}^2$ repeatedly to get

$$\mathcal{G}(Q) \leq \mathcal{G}^3(Q) \leq \mathcal{G}^5(Q) \leq \cdots \mathcal{G}^4(Q) \leq \mathcal{G}^2(Q) \leq Q.$$

Apply $\mathcal{G}^2$ repeatedly to get

$$\mathcal{G}(Q) \le \mathcal{G}^3(Q) \le \mathcal{G}^5(Q) \le \cdots \mathcal{G}^4(Q) \le \mathcal{G}^2(Q) \le Q.$$

The sequence $\mathcal{G}^{2k}(Q)$ decreases and has a lower bound. Hence there is a limit X_∞.

The sequence $\mathcal{G}^{2k+1}(Q)$ increases and has an upper bound. Hence there is a limit $X_{-\infty}$.

Moreover $X_{-\infty} \le X_\infty$

Moreover $X_{-\infty} \leq X_{\infty}$

- If $X_{-\infty} = X_{\infty}$ then it is the global attractor of $\mathcal{G}$: for every $X > 0$ for which $\mathcal{G}(X) > 0$ we have $\mathcal{G}^j(X) \to X_{\infty} = X_{-\infty}$.

Moreover $X_{-\infty} \leq X_{\infty}$

- If $X_{-\infty} = X_{\infty}$ then it is the global attractor of $\mathcal{G}$: for every $X > 0$ for which $\mathcal{G}(X) > 0$ we have $\mathcal{G}^j(X) \to X_{\infty} = X_{-\infty}$.

- If $X_{-\infty} \neq X_{\infty}$ then we have that trajectories starting from $X \leq X_{-\infty}$ converge to the periodic orbit $X_{-\infty}$, X_{∞}, trajectories starting with $X \geq X_{\infty}$ converge to that periodic orbit provided $\mathcal{G}(X) > 0$.

Example

$$X + A^*XA = Q$$

$\mathcal{G}(Q) > 0$ means $Q - A^*QA > 0$. As $Q > 0$ this implies A is stable w.r.t. the unit circle.

$$\mathcal{G}^2(X) = X$$

becomes

$$X = \mathcal{G}(Q) + A^{*2}XA^2$$

$$X - A^{*2}XA^2 = \mathcal{G}(Q)$$

Stein equation. A stable implies there is a unique solution. So

$$X_{-\infty} = X_{\infty}$$

and $X + A^*XA = Q$ has a unique solution.

$\mathcal{G}(Q) > 0$ is necessary here.

$\mathcal{G}(Q) > 0$ is necessary here.

Example:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \qquad Q = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$\mathcal{G}(Q) > 0$ is necessary here.

Example:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \qquad Q = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Then $A^2 = I$ and $\mathcal{G}(Q) = 0$.

So: any $X > 0$ satisfies

$$\mathcal{G}^2(X) = X$$

However, X solves

$$X + A^* X A = Q$$

if and only if

$$X = \begin{pmatrix} 1 & x \\ \bar{x} & \frac{1}{2} + \text{Re } x \end{pmatrix}$$

Thus a solution X is positive if and only if

$$|x - \frac{1}{2}| < \frac{\sqrt{3}}{2}.$$

Linear matrix equations

Linear matrix equations

Consider the linear matrix equations

$$X - A_1^* X A_1 - \cdots - A_m^* X A_m = Q,$$

and

$$X + A_1^* X A_1 + \cdots + A_m^* X A_m = Q,$$

where $Q = Q^*$ and A_i are $n \times n$ matrices.

Special case $m = 1$ in the first one: the Stein equation.

$$X - A_1^* X A_1 = Q,$$

Then: if $Q > 0$ there is a unique positive definite solution if and only if A is stable. The solution is given by

$$X = \sum_{i=0}^{\infty} A_1^{*i} Q A_1^i.$$

Kronecker product

The Kronecker product of two $n \times n$ matrices A, B denoted by $A \otimes B$, is defined as

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nn}B \end{bmatrix}$$

where $A = (a_{ij})$.

Theorem 1. *A matrix X is a solution of equation*

$$X - A_1^* X A_1 - \cdots - A_m^* X A_m = Q,$$

if and only if $x = \mathrm{vec}(X)$ is a solution of $Kx = q$, with $K = I_{n^2} - \sum_{j=1}^{m} A_j^T \otimes A_j^*$ and $q = \mathrm{vec}(Q)$. Consequently, the equation has a unique solution for any Q if and only if the matrix K is nonsingular.

Theorem 1. *A matrix X is a solution of equation*

$$X - A_1^* X A_1 - \cdots - A_m^* X A_m = Q,$$

if and only if $x = \text{vec}(X)$ is a solution of $Kx = q$, with $K = I_{n^2} - \sum_{j=1}^m A_j^T \otimes A_j^*$ and $q = \text{vec}(Q)$. Consequently, the equation has a unique solution for any Q if and only if the matrix K is nonsingular.

Similar result for

$$X + A_1^* X A_1 + \cdots + A_m^* X A_m = Q$$

and $L = I_{n^2} + \sum_{j=1}^m A_j^T \otimes A_j^*$

Lemma 1. *Assume there is a* $\tilde{Q} > 0$ *such that*

$$\tilde{Q} - \sum_{j=1}^{m} A_j^* \tilde{Q} A_j > 0.$$

Then

$$K = I_{n^2} - \sum_{j=1}^{m} A_j^T \otimes A_j^*$$

is invertible. So in this case the first equation has a unique solution for any $n \times n$ *matrix* Q.

Idea of proof: first assume $\tilde{Q} = I$, i.e.,

$$I - \sum_{j=1}^{m} A_j^* A_j > 0.$$

Assume also K is not invertible. Then there is an $n \times n$ matrix X such that

$$X - A_1^* X A_1 - \cdots - A_m^* X A_m = 0.$$

Put

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix}, \quad \hat{X} = \begin{bmatrix} X & 0 & \cdots & 0 \\ 0 & X & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & X \end{bmatrix}$$

Then $X = A^* \hat{X} A$. Now use a norm estimate.

$$\|X\| = \|A^* \hat{X} A\| \leq \|A\|^2 \|\hat{X}\| =$$

$$= \|A^* A\| \|X\| = \| \sum_{j=1}^{m} A_j^* A_j \| \|X\|,$$

which is a contradiction.

The general case can be reduced to this case.

$X - \sum_{i=1}^{m} A_i^* X A_i = Q$

Introduce the map

$$\mathcal{G}(X) = Q + \sum_{j=1}^{m} A_j^* X A_j.$$

Theorem 2. *Let Q be positive definite and assume there is a positive definite solution $\tilde{X}_0$ of the inequality*

$$X - A_1^* X A_1 - \ldots - A_m^* X A_m \geq Q.$$

Then the equation

$$X - A_1^* X A_1 - \ldots - A_m^* X A_m = Q$$

has a unique solution.

Moreover, this unique solution is positive definite and the sequence $\{\mathcal{G}^k(Q)\}_{k=0}^{\infty}$ *increases to this unique solution; the sequence* $\{\mathcal{G}^k(\tilde{X}_0)\}_{k=0}^{\infty}$ *decreases to this unique solution.*

Moreover, this unique solution is positive definite and the sequence $\{\mathcal{G}^k(Q)\}_{k=0}^{\infty}$ increases to this unique solution; the sequence $\{\mathcal{G}^k(\tilde{X}_0)\}_{k=0}^{\infty}$ decreases to this unique solution.

Finally, the unique solution is given by the formula

$$\bar{X} = \lim_{k\to\infty} \mathcal{G}^k(Q) =$$

$$= Q + \sum_{i=1}^{\infty} \sum_{j_1,\dots,j_i=1}^{m} A_{j_1}^* \cdot \ldots \cdot A_{j_i}^* Q A_{j_i} \cdot \ldots \cdot A_{j_1}.$$

Compare the Stein equation and the formula for the solution.

$$X + \sum_{i=1}^{m} A_i^* X A_i = Q$$

Introduce the map

$$\mathcal{H}(X) = Q - \sum_{j=1}^{m} A_j^* X A_j.$$

Then $\mathcal{H}$ is order reversing.

$X + \sum_{i=1}^{m} A_i^* X A_i = Q$

Introduce the map

$$\mathcal{H}(X) = Q - \sum_{j=1}^{m} A_j^* X A_j.$$

Then $\mathcal{H}$ is order reversing.

Theorem 3. *If Q is positive definite and satisfies $Q - \sum_{j=1}^{m} A_j^* Q A_j > 0$, then the equation has a unique positive definite solution given by*

$$\bar{X} = Q + \sum_{i=1}^{\infty} \sum_{j_1, \ldots, j_i = 1}^{m} (-1)^i A_{j_1}^* \cdot \ldots \cdot A_{j_i}^* Q A_{j_i} \cdot \ldots \cdot A_{j_1}.$$

Idea of the proof: to show that $X_{-\infty} = X_{\infty}$.

These are solutions to $\mathcal{H}^2(X) = X$, and that turns out to be a linear matrix equation of the type treated earlier.

One shows that the condition $Q - \sum_{j=1}^{m} A_j^* Q A_j > 0$ implies that the latter equation has a unique solution.

Sahnovic equation

From an application in interpolation theory appears the following matrix equation

$$X = Q - A^*(D - \tilde{X})^{-1}A$$

where

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix},$$

$$\tilde{X} = \begin{bmatrix} X & & \\ & \ddots & \\ & & X \end{bmatrix}.$$

It is assumed that $D < \tilde{Q}$.

Introduce

$$\mathcal{G}(X) = Q - A^*(D - \tilde{X})^{-1}A$$

Then $\mathcal{G}$ is order reversing, at least for matrices $X > Q$.

As $D < \tilde{Q}$ we have $\mathcal{G}(Q) > Q$. Also, $\mathcal{G}$ maps $[Q, \mathcal{G}(Q)]$ into itself.

Introduce

$$\mathcal{G}(X) = Q - A^*(D - \tilde{X})^{-1}A$$

Then $\mathcal{G}$ is order reversing, at least for matrices $X > Q$.

As $D < \tilde{Q}$ we have $\mathcal{G}(Q) > Q$. Also, $\mathcal{G}$ maps $[Q, \mathcal{G}(Q)]$ into itself.

So, again we have $\mathcal{G}^{2j}(Q)$ increases to a limit $X_{-\infty}$, while $\mathcal{G}^{2j+1}(Q)$ decreases to a limit X_{∞}, and these two matrices form a periodic orbit of period two, or they coincide.

Introduce

$$\mathcal{G}(X) = Q - A^*(D - \tilde{X})^{-1}A$$

Then $\mathcal{G}$ is order reversing, at least for matrices $X > Q$.

As $D < \tilde{Q}$ we have $\mathcal{G}(Q) > Q$. Also, $\mathcal{G}$ maps $[Q, \mathcal{G}(Q)]$ into itself.

So, again we have $\mathcal{G}^{2j}(Q)$ increases to a limit $X_{-\infty}$, while $\mathcal{G}^{2j+1}(Q)$ decreases to a limit X_{∞}, and these two matrices form a periodic orbit of period two, or they coincide.

Numerical evidence suggest: always

$$X_{-\infty} = X_{\infty}.$$

Theorem 4. *If $D < \tilde{Q}$ then the equation*

$$\mathcal{G}(X) = Q - A^*(D - \tilde{X})^{-1}A$$

has a unique positive definite solution.

Remarks The theorem does not state that there is a unique solution, in fact it can have many.

Remarks The theorem does not state that there is a unique solution, in fact it can have many.

Trying to solve the equation with Newton's method one gets to solve linear matrix equations of the type treated earlier. Newton's method started at Q will occasionally converge to a solution that is not positive definite (i.e., not the largest solution).

Remarks The theorem does not state that there is a unique solution, in fact it can have many.

Trying to solve the equation with Newton's method one gets to solve linear matrix equations of the type treated earlier. Newton's method started at Q will occasionally converge to a solution that is not positive definite (i.e., not the largest solution).

There is still no good insight into the structure of the full set of solutions.

Idea of proof.

$$\mathcal{G}(X_\infty) = X_{-\infty}, \quad \mathcal{G}(X_{-\infty}) = X_\infty.$$

Consider

$$\begin{aligned} & X_\infty - X_{-\infty} = \\ = \quad & \mathcal{G}(X_{-\infty}) - \mathcal{G}(X_\infty) = \\ = \quad & A^*\{(D - \tilde{X}_\infty)^{-1} - (D - \tilde{X}_{-\infty})^{-1} A. \end{aligned}$$

Now

$$\begin{aligned} & (D - \tilde{X}_\infty)^{-1} - (D - \tilde{X}_{-\infty})^{-1} = \\ = \quad & (D - \tilde{X}_\infty)^{-1}(\tilde{X}_\infty - \tilde{X}_{-\infty})(D - \tilde{X}_{-\infty})^{-1}. \end{aligned}$$

So $Y = X_\infty - X_{-\infty}$ satisfies the linear matrix equation

$$Y = \sum_{j=1}^{m} B_j Y C_j$$

where

$$B = A^*(D - \tilde{X}_\infty)^{-1}, \quad C = (D - \tilde{X}_{-\infty})^{-1} A.$$

Unfortunately, this is not a symmetric linear matrix equation. However, a symmetric version of this can be derived in a similar way. And for that one the condition in the theorem just tells us that the linear equation has a unique solution. As $Y = 0$ solves the linear equation in question we get $X_\infty = X_{-\infty}$.