מערכות דינאמיות ובקרה

לביא שפיגלמן

30 בנובמבר 2006

- פירוק מטריצה ע"י SVD
- שיערוך תוך מיזעור שגיאה ריבועית (שערוך מצב, system identification)
  - בעיה over constrained
  - בעיה under constrained
  - מציאת אות בקרה אופטימלי
  - שערוך מצב עם רעש תצפית
  - שערוך מודל

Singular Value Decomposition

תהי $\mathbf{A} \in \mathrm{R}^{m \times n}$ שדרגתה r. אם $m \geq n$ (או $m = n = k$) אך לא קיימים k וקטורים עצמיים בלתי תלויים של $\mathbf{A}$) אז לא ניתן ללכסן את $\mathbf{A}$. כלומר, אין מטריצה $\mathbf{P}$ כך ש$\mathbf{A} = \mathbf{PDP}^{-1}$ ו$\mathbf{D}$ אלכסונית. לעומת זאת, תמיד קיים Singular Value Decomposition:

$$\mathbf{A} = \mathbf{U\Sigma V}'$$

כך ש$\mathbf{U} \in \mathrm{R}^{m \times r}$, $\mathbf{V} \in \mathrm{R}^{n \times r}$, $\mathbf{U'U} = \mathbf{V'V} = \mathbf{I}$ (אורטונורמליות) ו$\mathbf{\Sigma} \in \mathrm{R}^{r \times r}$ אלכסונית,ואיברי האלכסון אינם 0.

הערה: מכאן גם ש$\mathbf{\Sigma}^{-1}$ קיימת ואם נגדיר $\mathbf{B} = \mathbf{V\Sigma}^{-1}\mathbf{U}'$ אז קל לראות ש $\mathbf{BA} = \mathbf{I}_n$. לכן $\mathbf{B}$ הוא מעין הופכי (שמאלי) של $\mathbf{A}$.

הערה: $\mathbf{A'A} = \mathbf{V\Sigma U'U\Sigma V'} = \mathbf{V\Sigma}^2\mathbf{V}'$ ובאופן דומה $\mathbf{AA}' = \mathbf{U\Sigma}^2\mathbf{U}'$. מכאן ש $Rank(\mathbf{A'A}) = Rank(\mathbf{AA}') = r$. אם $r = n$ אז $\mathbf{A'A} \in \mathrm{R}^{n \times n}$ הפיכה ו$(\mathbf{A'A})^{-1} = \mathbf{V\Sigma}^{-2}\mathbf{V}'$
אם $r = m$ אז $\mathbf{AA}' \in \mathrm{R}^{m \times m}$ הפיכה ו$(\mathbf{AA}')^{-1} = \mathbf{U\Sigma}^{-2}\mathbf{U}'$.

שיערוך תוך מזעור שגיאה ריבועית

בעיה 1: (**over constrained**)

נניח כי נתונים $\mathbf{y} \in \mathrm{R}^q$ ו$\mathbf{A} \in \mathrm{R}^{q \times n}$ כך ש $q > n$ ($\mathbf{A}$ רזה) וכי $Rank(A) = n$

ברצוננו למצוא $\mathbf{x} \in \mathrm{R}^n$ כך ש-$\mathbf{Ax}$ קרוב ל-$\mathbf{y}$ שממזער את השגיאה הריבועית:

$$\arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{Ax}\|^2$$

זה ביטוי ריבועי ב$\mathbf{x}$ ולכן נקודה בה הנגזרת מתאפסת היא נקודת מינימום. על מנת להקל את הגזירה נפתח מעט את הביטוי:

$$J = \|\mathbf{y} - \mathbf{Ax}\|^2 = (\mathbf{y} - \mathbf{Ax})'(\mathbf{y} - \mathbf{Ax}) = \mathbf{y}'\mathbf{y} - 2\mathbf{y}'\mathbf{Ax} + \mathbf{x}'\mathbf{A}'\mathbf{Ax}$$

נגזור לפי $\mathbf{x}$ ונשווה ל0

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{x}} = -2\mathbf{y}'\mathbf{A} + 2\mathbf{x}'\mathbf{A}'\mathbf{A} &= 0 \\ \mathbf{A}'\mathbf{y} &= \mathbf{A}'\mathbf{Ax} \end{aligned}$$

כעת נעזר בכך ש$Rank(\mathbf{A}) = n$ ולכן $\mathbf{A}'\mathbf{A}$ הפיכה. נכפול משמאל ב $(\mathbf{A}'\mathbf{A})^{-1}$ לקבלת

$$\mathbf{x} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{y}$$

אם נציב $\mathbf{A} = \mathbf{U\Sigma V}'$ נקבל $\mathbf{x} = \mathbf{V\Sigma}^{-1}\mathbf{U}'\mathbf{y}$. הביטוי $(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A} = \mathbf{V\Sigma}^{-1}\mathbf{U}'$ ממזער את השגיאה הריבועית ונקרא הופכי שמאלי של $\mathbf{A}$ (כי כפל ב$\mathbf{A}$ מימינו נותן $\mathbf{I}$). נסמנו ב$\mathbf{A}^\dagger$ (ההופכי הימני שיוצג בהמשך מסומן באופן זהה)

בעיה 2: (**under constrained**)

נניח כי נתונים $\mathbf{y} \in \mathrm{R}^q$ ו$\mathbf{A} \in \mathrm{R}^{q \times n}$ כך ש $q < n$ ($\mathbf{A}$ שמנה) וכי $Rank(A) = n$

ברצוננו למצוא $\mathbf{x} \in \mathrm{R}^n$ כך ש-$\mathbf{Ax} = \mathbf{y}$ (ניתו לעשות זאת) ו-$\mathbf{x}$ קטן ביותר (במובן של נורמה ריבועית):

$$\arg\min_{x} ||\mathbf{x}||^2 \qquad s.t.\ \mathbf{Ax} = \mathbf{y}$$

זו בעיה עם אילוצים. נפתור בעזרת כופלי לגרנג':

$$J = \mathbf{x}'\mathbf{x} + \lambda'(\mathbf{Ax} - \mathbf{y})$$

נגזור לפי $\mathbf{x}$

$$\begin{aligned} \frac{\partial J}{\partial x} &= 2\mathbf{x}' + \lambda'\mathbf{A} = 0 \\ \mathbf{x} &= \frac{1}{2}\mathbf{A}'\lambda \end{aligned}$$

נמצא λ שמקיים את האילוץ

$$\mathbf{A}(\frac{1}{2}\mathbf{A}'\lambda) = \mathbf{y}$$

וכיוון שקיים $(\mathbf{A}\mathbf{A}')^{-1}$

$$\lambda \quad = \quad 2(\mathbf{A}\mathbf{A}')^{-1}\mathbf{y}$$

ונציב חזרה ב$\mathbf{x}$

$$\mathbf{x} = \mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}\mathbf{y} = \mathbf{A}^{\dagger}\mathbf{y}$$

לקבלת הpseudo inverse הימני. נשים לב שהצבת $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}'$ נותנת כמיקודם $\mathbf{A}^{\dagger} = \mathbf{V}\Sigma^{-1}\mathbf{U}$

מציאת אות בקרה אופטימלי

תהי מערכת LTI:

$$\mathbf{x}_{n+1} \quad = \quad \mathbf{A}\mathbf{x}_n + \mathbf{B}\mathbf{u}_n \quad , \mathbf{x}_0 = \mathbf{0} \;\; , \;\; \mathbf{u}_0 = \mathbf{0}$$

נרצה למצוא אות בקרה שיביא את המערכת בצעד הn למצב נתון $\mathbf{x}_n$.

כידוע, ניתן לפרוש את נוסחת הנסיגה לקבלת

$$\mathbf{x}_n = \sum_{i=0}^{n-1} \mathbf{A}^i \mathbf{B}\mathbf{u}_i$$

ובאופן מטריציוני

$$\mathbf{x}_n = \underbrace{\begin{bmatrix}\mathbf{B} & \mathbf{A}\mathbf{B} & \cdots & \mathbf{A}^{n-1}\mathbf{B}\end{bmatrix}}_{\tilde{\mathbf{A}}} \underbrace{\begin{bmatrix}\mathbf{u}_0 \\ \vdots \\ \mathbf{u}_{n-1}\end{bmatrix}}_{\tilde{\mathbf{u}}}$$

זו בעיה under constrained ולכן, אם $\tilde{\mathbf{A}}$ בעלת דרגת שורות מלאה (היינו, אם המערכת ניתנת לבקרה!) אז ישנם הרבה $\tilde{\mathbf{u}}$ שפותרים את המשוואה.

אם נבחר את $\tilde{\mathbf{u}}$ כך ש $\|\tilde{\mathbf{u}}\|^2$ מינימלי הרי שהפתרון הוא, כמיקודם, $\tilde{\mathbf{u}} = \tilde{\mathbf{A}}'(\tilde{\mathbf{A}}\tilde{\mathbf{A}}')^{-1}\mathbf{x}_n$.
אם נציב בחזרה את $\tilde{\mathbf{A}}$ ונבצע את מכפלות המטריצות נקבל

$$\mathbf{u}_i = \mathbf{B}'\mathbf{A}^i \underbrace{\left(\sum_{j=0}^{n-1} \mathbf{A}^i \mathbf{B}\mathbf{B}' \left(\mathbf{A}^i\right)'\right)^{-1}}_{\mathbf{W}_c^{-1}(n-1)} \mathbf{x}_n$$

נשים לב שזהו בדיוק אות הבקרה ששימש בהוכחה כי אם ה- $\text{controllability grammian}$ הפיכה אז המערכת ניתנת לבקרה (עבור מערכת דיסקרטית). גם כאן נדרשנו להנחה זו (איפו?).

שערוך מצב עם רעש תצפית

נניח כי נתונה לנו מערכת (דיסקרטית בזמן)

$$\begin{aligned}\mathbf{x}_t &= \mathbf{A}\mathbf{x}_{t-1} \\ \mathbf{y}_t &= \mathbf{C}\mathbf{x}_t + \nu_t\end{aligned}$$

כש$\nu_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ רעש נורמלי לבן. נניח כי המערכת היא observable ונניח כי ראינו את סדרת התצפיות $\{\mathbf{y}_0, \dots \mathbf{y}_t\}$. נשים לב כי

$$\underbrace{\begin{bmatrix}\mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_t\end{bmatrix}}_{\bar{\mathbf{y}}} = \underbrace{\begin{bmatrix}\mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{t-1}\end{bmatrix}}_{\mathcal{O}} \mathbf{x}_0 + \begin{bmatrix}\nu_0 \\ \nu_1 \\ \vdots \\ \nu_t\end{bmatrix}$$

ולכן

$$\mathbf{y}_i \sim \mathcal{N}(\mathbf{C}\mathbf{A}^i\mathbf{x}_0, \mathbf{I})$$

נחפש

$$\arg\max_{\hat{\mathbf{x}}_0} p(\mathbf{y}_0, \dots \mathbf{y}_t|\hat{\mathbf{x}}_0) = \arg\max_{x_0} \prod_{i=0}^{t} \frac{1}{z} e^{-\frac{1}{2}||\mathbf{y}-\mathbf{C}\mathbf{A}^i\hat{\mathbf{x}}_0||^2}$$

כידוע הפעולה $\ln(x)$ היא מונוטונית ועולה בx ולכן $\arg\max_x f(x) = \arg\max_x \ln(f(x))$. לכן די למצוא

$$\begin{aligned}\arg\max_{\hat{\mathbf{x}}_0} \ln(p(\mathbf{y}_0, \dots \mathbf{y}_t|\hat{\mathbf{x}}_0)) &= \arg\max_{\hat{\mathbf{x}}_0} \sum_{i=0}^{t} -\frac{1}{2}||\mathbf{y} - \mathbf{C}\mathbf{A}^i\hat{\mathbf{x}}_0||^2 \\ &= \arg\min_{\hat{\mathbf{x}}_0} \frac{1}{2}||\tilde{\mathbf{y}} - \mathcal{O}\mathbf{x}_0||^2\end{aligned}$$

כלומר, מההנחה של רעש נורמלי קיבלנו שפתרון הניראות המירבית הוא פתרון בעל שגיאה ריבועית מינימלית. כפי שראינו

$$\begin{aligned}\hat{\mathbf{x}}_0 &= (\mathcal{O}'\mathcal{O})^{-1}\mathcal{O}\bar{\mathbf{y}} \\ &= \left(\sum_{k=0}^{t-1} \mathbf{A}'^k\mathbf{C}'\mathbf{C}\mathbf{A}^\mathbf{k}\right)^{-1} \sum_{k=0}^{t-1} \mathbf{A}'^k\mathbf{C}'\mathbf{y}_k\end{aligned}$$

הנחת הניצפות מבטיחה קיום $(\mathcal{O}'\mathcal{O})^{-1}$.

כעת נניח כי $|\nu_k| < v$ (לא בהכרח גאוסיאני) ונבחן מה קורה לשגיאה בשערוך $\mathbf{x}_0$ כתלות במשך המדידה t:

$$\begin{aligned}\hat{\mathbf{x}}_0 - \mathbf{x}_0 &= \mathcal{O}^\dagger\bar{\mathbf{y}} - \mathcal{O}^\dagger(\bar{\mathbf{y}} - \bar{\nu}) \\ &= \mathcal{O}^\dagger\bar{\nu} \\ &= \left(\sum_{k=0}^{t-1} \mathbf{A}'^k\mathbf{C}'\mathbf{C}\mathbf{A}^\mathbf{k}\right)^{-1} \sum_{k=0}^{t-1} \mathbf{A}'^k\mathbf{C}'\nu_k\end{aligned}$$

למען פשטות הניתוח נניח כי $\mathbf{C} = 1$ ו$\mathbf{A} = a$ לקבלת

$$|\hat{\mathbf{x}}_0 - \mathbf{x}_0| = \frac{\sum_{k=0}^{t-1} a^t \nu_k}{\sum_{k=0}^{t-1} a^{2t}} \leq \frac{v \sum_{k=0}^{t-1} a^t}{\sum_{k=0}^{t-1} a^{2t}}$$

זה סכום של סדרות הנדסיות. נזכור כי $\sum_{k=0}^{t} a^k = \frac{1-a^{k+1}}{1-a}$.
אם $|a| < 1$ כש$t \to \infty$ נקבל

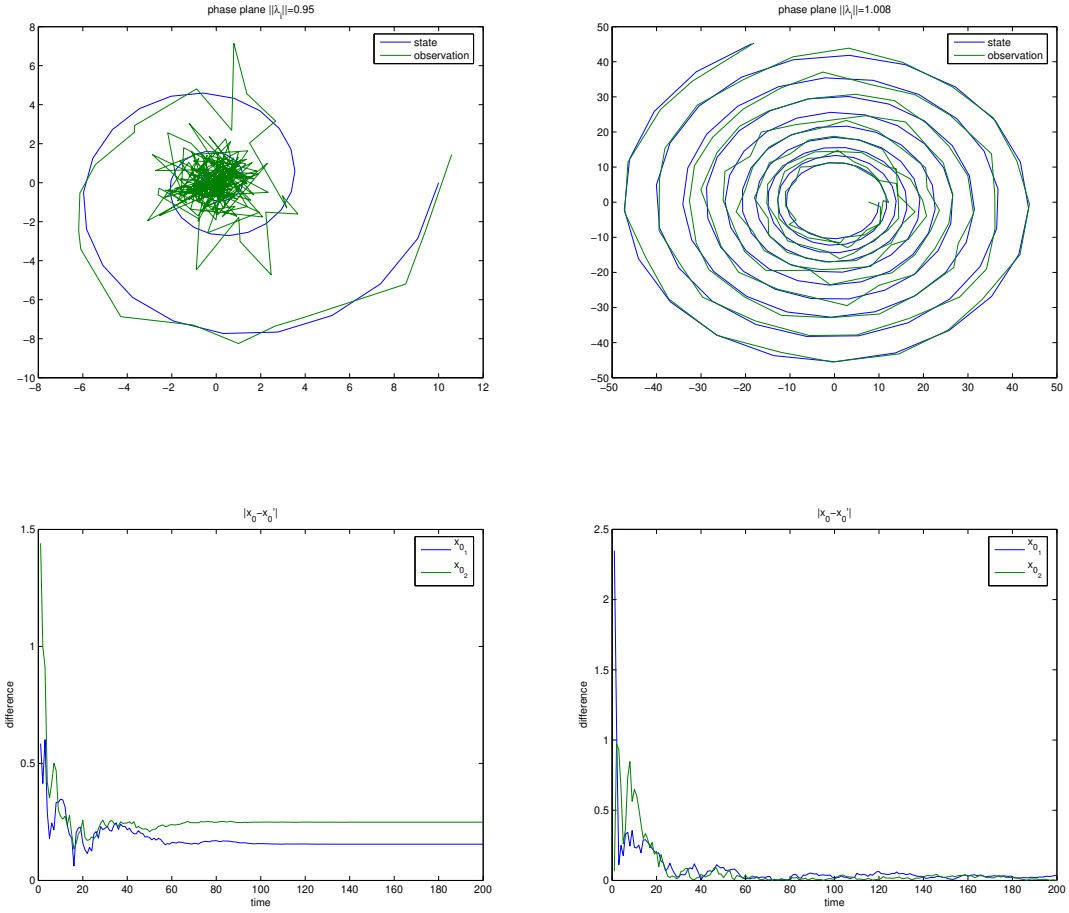
$$\hat{\mathbf{x}}_0 - \mathbf{x}_0 \leq \frac{v(1-a^2)}{(1-a)} = constant$$

אחרת

$$\hat{\mathbf{x}}_0 - \mathbf{x}_0 = 0$$

אם $\mathbf{A}$ יציבה אז המצב ההתחלתי 'נשכח' עם הזמן ולכן תצפיות נוספות מפסיקות לשפר את השערוך. אם $\mathbf{A}$ לא יציבה אז סטייה קטנה בערך ההתחלתי גדלה בקצב אקספוננציאלי ולכן נוכל להקטין את השגיעה ל0.

דוגמא:



בשורה העליונה: מסלולים של המצב והתצפיות. בשורה התחתונה: $|\hat{\mathbf{x}}_0 - \mathbf{x}_0|$ כתלות במשך התצפית. משמאל: עבור מערכת יציבה. מימין - עבור מערכת מתבדרת.

ראינו שהשגיאה היא $\mathcal{O}^{\dagger}\bar{\nu}$ מכאן שהיינו רוצים pseudo inverse 'קטן' ביותר. נראה כי $\mathcal{O}^{\dagger}$ מקיים זאת.

טענה: תהי $\mathbf{A} \in \mathrm{R}^{q\times n}$ רזה בעלת דרגה מלאה ו$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}'$. תהי $\mathbf{A}^{\dagger} = \mathbf{V}\Sigma^{-1}\mathbf{U}'$. לכל מטריצה $\mathbf{B}$ כך ש $\mathbf{BA} = \mathbf{I}$ מתקיים $\mathbf{BB}' - \mathbf{A}^{\dagger}(\mathbf{A}^{\dagger})' \succeq \mathbf{0}$

($\mathbf{M} \succeq \mathbf{0}$ מסמל ש $\mathbf{M}$ היא positive semi definite, כלומר לכל וקטור x מתקיים ש $x'\mathbf{M}x > 0$)

משמעות: לכל וקטור (רעש) ν מתקיים ש $||\mathbf{A}^{\dagger}\nu||^2 < ||\mathbf{B}\nu||^2$

הוכחה:

נגדיר $\mathbf{Z} = \mathbf{B} - \mathbf{A}^\dagger$, כלומר $\mathbf{B} = \mathbf{A}^\dagger + \mathbf{Z}$

קל לראות כי

$$\mathbf{ZA} = (\mathbf{B} - \mathbf{A}^\dagger)\mathbf{A} = \mathbf{I} - \mathbf{I} = 0$$

לכן

$$\begin{aligned} 0 &= \mathbf{ZA} = \mathbf{ZU\Sigma V}' \\ 0\mathbf{V}\Sigma^{-1} &= \mathbf{ZU\Sigma V}'\mathbf{V}\Sigma^{-1} \\ 0 &= \mathbf{ZU} \end{aligned}$$

מכאן ש

$$\mathbf{Z}(\mathbf{A}^\dagger)' = (\mathbf{ZU})\Sigma^{-1}\mathbf{V}' = 0$$

ולכן

$$\begin{aligned} \mathbf{BB}' &= (\mathbf{A}^\dagger + \mathbf{Z})(\mathbf{A}^\dagger + \mathbf{Z})' \\ &= \mathbf{A}^\dagger(\mathbf{A}^\dagger)' + \mathbf{A}^\dagger\mathbf{Z}' + (\mathbf{A}^\dagger\mathbf{Z}')' + \mathbf{ZZ}' \\ &= \mathbf{A}^\dagger(\mathbf{A}^\dagger)' + \mathbf{ZZ}' \\ \mathbf{BB}' - \mathbf{A}^\dagger(\mathbf{A}^\dagger)' &\succeq \mathbf{0} \end{aligned}$$

מערכת זמן רציף:

נתונה מערכת בזמן רציף

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{Ax}(t) \\ \mathbf{y}(t) &= \mathbf{Cx}(t) + \nu(t) \end{aligned}$$

השגיעה הריבועית בין תצפיות (אמיתית ומשוערכת) היא

$$\begin{aligned} J &= \int_0^t ||\mathbf{y}(\tau) - \mathbf{C}e^{\mathbf{A}\tau}\hat{\mathbf{x}}(0)||^2 d\tau \\ &= \mathbf{x}(0)'\mathbf{Qx}(0) + 2\mathbf{r}'\mathbf{x}(0) + s \\ \mathbf{Q} &= \mathbf{W}_o(t) = \int_0^t e^{\mathbf{A}'\tau}\mathbf{C}'\mathbf{C}e^{A\tau}d\tau \\ \mathbf{r} &= \int_0^t e^{\mathbf{A}'\tau}\mathbf{C}'\mathbf{y}(\tau)d\tau \\ s &= \int_0^t \mathbf{y}'(\tau)\mathbf{y}(\tau)d\tau \end{aligned}$$

באופן דומה ניתן לגזור את השגיאה לקבלת $\hat{\mathbf{x}}(0) = \mathbf{Q}^{-1}\mathbf{r}$ ושגיאת השערוך היא

$$\hat{\mathbf{x}}(0) - \mathbf{x}(0) = \mathbf{Q}^{-1}\int_0^t e^{\mathbf{A}'\tau}\mathbf{C}'\nu(\tau)d\tau$$

System Identification

נניח כי נתונות דוגמאות $\{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^N$ וברצוננו למצוא $\mathbf{A}$ כך ש$\mathbf{Ax}$ קרוב ל$\mathbf{y}$. נגדיר בעיית שערוך

$$\arg\min_{\mathbf{A}} \sum_{i=1}^{N} ||\mathbf{y}_i - \mathbf{A}\mathbf{x}_i||^2 = \sum_{i=1}^{N} \mathbf{x}_i'\mathbf{A}'\mathbf{A}\mathbf{x}_i - 2\mathbf{y}_i'\mathbf{A}\mathbf{x}_i + \mathbf{y}_i'\mathbf{y}_i$$

נגזור לפי $\mathbf{A}$ לקבלת

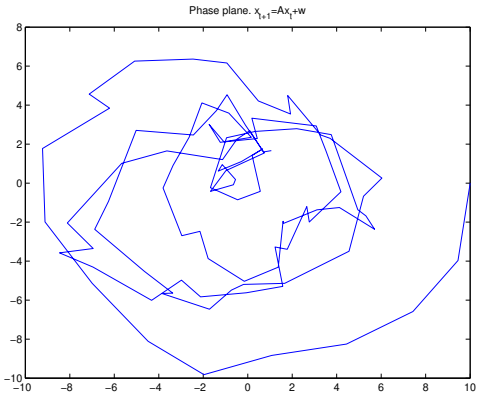
$$\begin{aligned}
\frac{\partial}{\partial \mathbf{A}} \sum_{i=1}^{N} ||\mathbf{y}_i - \mathbf{A}\mathbf{x}_i||^2 &= 0 \\
\sum_{i=1}^{N} 2\mathbf{A}\mathbf{x}_i\mathbf{x}_i' - 2\mathbf{y}_i\mathbf{x}_i' &= 0 \\
\mathbf{A}\sum_{i=1}^{N} \mathbf{x}_i\mathbf{x}_i' &= \sum_{i=1}^{N} \mathbf{y}_i\mathbf{x}_i'
\end{aligned}$$

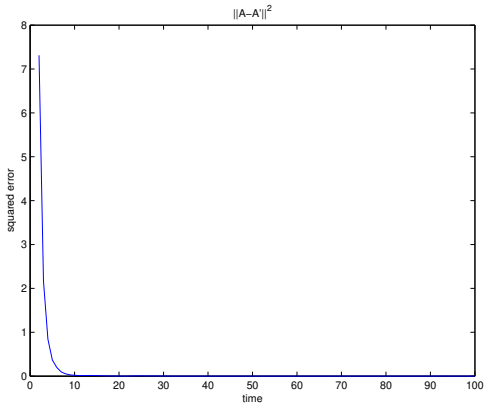
נניח שדרגת $\sum_{i=1}^{N} \mathbf{x}_i\mathbf{x}_i'$ מלאה (תנאי הכרחי הוא ש$N > n$) ונקבל

$$\mathbf{A} = \left(\sum_{i=1}^{N} \mathbf{y}_i\mathbf{x}_i'\right)\left(\sum_{i=1}^{N} \mathbf{x}_i\mathbf{x}_i'\right)^{-1}$$

דוגמא:

נרצה לשערך את $\mathbf{A}$ במערכת $\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \omega$





משמאל מסלול שהתקבל. מימין $||\mathbf{A} - \hat{\mathbf{A}}||^2$ כתלות במספר הדגימות.