

Linjär algebra V 2011.

LAB 2: Felanalys vid lösning av ekvationssystem.

Inledning

När man behandlar ett ekvationssystem

$$A\mathbf{x} = \mathbf{b} \quad (1)$$

är ofta matrisen A och högerledet $\mathbf{b}$ behäftade med fel eller osäkerheter. De data som matas in kan t.ex. vara resultat av mätningar som har en begränsad noggrannhet. Dessutom kan oftast inte talen representeras exakt i en dator, vilket ger små fel som dock kan förstöras av ett stort antal räkneoperationer. Det system som vi i praktiken arbetar med har då matrisen $A + \delta A$ och högerledet $\mathbf{b} + \delta \mathbf{b}$ istället för A respektive $\mathbf{b}$, där δA och $\delta \mathbf{b}$ betecknar en *störning* av A respektive $\mathbf{b}$. Den lösning vi får kan skrivas $\mathbf{x} + \delta \mathbf{x}$, där $\mathbf{x}$ är lösningen till (1). Vi skall se hur man kan uppskatta storleken av felet $\delta \mathbf{x}$.

Ett naturligt mått på en vektor är dess längd, den *euklidiska normen*

$$\|\mathbf{v}\| = \sqrt{\sum_{i=1}^n v_i^2}.$$

Den euklidiska normen av en matris definieras som

$$\|A\| = \max_{\mathbf{x} \in \mathbb{R}^n} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}. \quad (2)$$

Ekvivalent är $\|A\|$ det minsta tal så att uppskattningen

$$\|A\mathbf{x}\| \leq \|A\| \|\mathbf{x}\| \quad (3)$$

gäller för alla $\mathbf{x}$.

Låt oss först betrakta situationen då vi har en störning enbart i högerledet, inte i matrisen. Vi antar också att A är *inverterbar*. Vi vill alltså lösa

$$A\mathbf{x} = \mathbf{b} \quad (4)$$

men behandlar i själva verket

$$A(\mathbf{x} + \delta \mathbf{x}) = \mathbf{b} + \delta \mathbf{b}.$$

Härav får vi att felet i lösningen uppfyller

$$\begin{aligned} A\delta \mathbf{x} &= \delta \mathbf{b}, \\ \delta \mathbf{x} &= A^{-1}\delta \mathbf{b}. \end{aligned} \quad (5)$$

Nu vill vi jämföra storleken av störningen i högerledet med storleken av felet i lösningen, eller mer precist sambandet mellan de relativa felen

$$\frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|}, \quad \frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|}.$$

Använder vi (3) på likheterna (4) och (5) får vi

$$\|\mathbf{b}\| \leq \|A\| \|\mathbf{x}\|, \quad \|\delta \mathbf{x}\| \leq \|A^{-1}\| \|\delta \mathbf{b}\|.$$

Kombinerar vi dessa olikheter får vi

$$\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \|A\| \|A^{-1}\| \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|}.$$

Produkten $\kappa(A) = \|A\| \|A^{-1}\|$ kallas *konditionstalet* för matrisen A , och man skriver

$$\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa(A) \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|}. \quad (6)$$

En matris med stort konditionstal sägs vara *illakonditionerad*. För sådana matriser kan alltså små störningar i högerledet ge upphov till förhållandevis stora fel i lösningen.

Om vi istället har en störning δA i matrisen med motsvarande fel $\delta \mathbf{x}$ i lösningen, kan vi på liknande sätt härleda feluppskattningen

$$\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x} + \delta \mathbf{x}\|} \leq \kappa(A) \frac{\|\delta A\|}{\|A\|}. \quad (7)$$

Vänsterledet är här något annorlunda, men vi ser att konditionstalet fortfarande avgör hur känslig ekvations-systemets lösning är för störningar.

Laborationsuppgift

Hilbertmatrisen H_n av ordning $n \times n$ har elementen $h_{i,j} = \frac{1}{i+j-1}$ och är ett exempel på en illakonditionerad matris (stort konditionstal). Hilbertmatrisen av ordning 3 är alltså

$$H_3 = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}$$

MATLAB har en funktion **hilb**(n) som genererar denna matris och en funktion **invhilb**(n) som genererar inversen med stor noggrannhet - dess element är heltal! Därför lämpar sig hilbertmatrisen väl för vår undersökning. För några n , titta på H_n och H_n^{-1} samt konditionstalet $\kappa(H_n)$. Konditionstalet ges i MATLAB av kommandot **cond**. Matrisnormen ges av kommandot **norm**.

1. Rita en graf som visar hur konditionstalet $\kappa(H_n)$ växer med n . Använd **semilogy** (linlog-skala: linjär i x-led, logaritmisk i y-led). Vad händer om vi samtidigt plottar $\kappa(H_n)$ och $\kappa(H_n^{-1})$ för $n \leq 20$? Vad kan det bero på?
2. För att förstå att vi har att göra med en lömsk matris, ska vi göra ett litet experiment som demonstrerar dess känslighet. Tag en Hilbertmatris av måttligt stor ordning, t ex $n = 4$. Välj ett högerled $\mathbf{b}$ och lös systemet $H_n \mathbf{x} = \mathbf{b}$ genom att använda kommandot **rref**. Nu "stör" vi matrisen H_n genom att runda av alla dess element till 4 decimaler (kommandot **round** kan utnyttjas). Lös systemet med samma högerled med **rref** och jämför resultatet. Kontrollera att resultatet överensstämmer med feluppskattningen (7)!
3. Här ska vi se hur beräkningarna i lösningsalgoritmen påverkar resultatet. Vi löser ekvationssystemet $H_n \mathbf{x} = \mathbf{b}$ för något värde på n , t.ex. $n = 10$ (ej större!). Tag $\mathbf{b}$ som en slumpvectorsvektor (**randn** i MATLAB: kommandot $b = \text{randn}(n, 1)$ genererar en kolonnvektor med n normalfördelade slumpstal). Jämför de lösningar du får med **rref**, $H_n \backslash \mathbf{b}$, $\text{inv}(H_n) * \mathbf{b}$ och den exakta lösningen **invhilb**(n) * $\mathbf{b}$.
4. Gör en funktionsfil som för Hilbertmatrisen av ordning n löser systemet $H_n \mathbf{x} = \mathbf{b}$ exakt för ett stort antal, N , slumpvis valda högerled $\mathbf{b}$. Tag till exempel $n = 3$, $N = 10^4$. Tag också N slumpvis valda felvektorer $\delta \mathbf{b}$. Motsvarande fel i lösningen kan beräknas exakt som $\delta \mathbf{x} = \text{invhilb}(\mathbf{n}) * \delta \mathbf{b}$; se (5). Beräkna sedan förhållandet mellan de relativa felen i lösningen och högerledet, det vill säga

$$\frac{\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|}}{\frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|}}.$$

För att få en bild av hur dessa kvoter varierar, spara dem i en vektor, sortera elementen i storleksordning (kommandot **sort**), och plotta resultatet. Jämför med feluppskattningen (6). Ligger den största kvoten i närheten av konditionstalet?

5. Upprepa uppgift 4 men undersök nu vad som händer då n och N varierar. Vilka slutsatser kan du dra?