NMB - Oefenzitting 2:
Conditie van kleinste kwadratenprobleem en stabiliteit van kleinste kwadratenalgoritmes

Hendrik Speleers

1 Pen en papier

Opgave 1. In het handboek wordt bij het onderzoeken van de conditie gebruik gemaakt van de parameter $\eta = \frac{\|A\|\|x\|}{\|y\|} = \frac{\|A\|\|x\|}{\|Ax\|}$. Men stelt daar $1 \leq \eta \leq \kappa(A)$. Toon dit aan. Voor welke x wordt de ondergrens (bovengrens) bereikt?

Opgave 2. Zij $A \in \mathbb{R}^{m\times n}$ met $m > n$ en rang n. De matrix A heeft een singuliere waardenontbinding $A = U\Sigma V^T$. De matrix U_1 bevat de eerste n kolommen van U, de matrix U_2 de volgende $m-n$. Bewijs dat $\text{range}(A) = \text{range}(U_1)$.

Opgave 3. In de cursus komen verschillende algoritmen aan bod om een kleinste kwadratenbenadering te berekenen. De keuze van een bepaald algoritme voor een bepaald probleem is niet altijd even eenvoudig. Zoek met behulp van de resultaten van les 10, 18 en 19 (Trefethen & Bau) een antwoord op volgende vraagjes. (Hou ook rekening met de rekencomplexiteit van elk van de algoritmen.)

a) Wat veroorzaakt het verschil in complexiteit tussen Householder triangularisatie en gewijzigde Gram-Schmidt?

b) Zou je eerder Householder triangularisatie of gewijzigde Gram-Schmidt orthogonalisatie gebruiken om een QR-factorisatie te bepalen van de coëfficiëntenmatrix A?

c) Zou je van mening veranderen indien je zou werken met de uitgebreide matrix $\begin{bmatrix} A & b \end{bmatrix}$?

d) Welke algoritmes zou je wel of niet gebruiken in de volgende gevallen en waarom?

- de matrix A is goed geconditioneerd, bv. $\kappa(A) \approx 1$
- de oplossingsvector b ligt dicht bij de kolomruimte van A en A is slecht geconditioneerd
- de oplossingsvector ligt niet dicht bij de kolomruimte en bovendien is je matrix slecht geconditioneerd

2 Op de computer

2.1 Conditie

We gaan het theoretische resultaat

$$\frac{\|\delta x\|}{\|x\|} \leq \kappa_{b\to x}\frac{\|\delta b\|}{\|b\|} \quad \text{met} \quad \kappa_{b\to x} = \frac{\kappa(A)}{\eta\cos(\theta)}$$

verifiëren aan de hand van experimenten.

We werken met een overgedetermineerd stelsel waarvan we de oplossing goed kennen. Hiervoor gaan we als volgt te werk. Kies een diagonaalmatrix $\Sigma \in \mathbb{R}^{10\times 7}$ en een vector $c \in \mathbb{R}^{10}$. Het overgedetermineerde stelsel

$$\Sigma z = c \qquad \text{met} \qquad z \in \mathbb{R}^7$$

kan opgelost worden met een relatieve fout van de grootte-orde van de machine-nauwkeurigheid. Kies $U \in \mathbb{R}^{10\times 10}$ en $V \in \mathbb{R}^{7\times 7}$ twee willekeurige orthogonale matrices en stel

$$A = U\Sigma V^T, \qquad b = Uc, \qquad x = Vz.$$

Uiteraard maakt men afrondingsfouten bij de numerieke berekening van A, b en x, maar doordat vermenigvuldigen met een orthogonale matrix goed geconditioneerd is, worden deze afrondingsfouten niet opgeblazen en mag men zeggen dat A, b en x berekend zijn met een relatieve fout van de grootte-orde van de machine-nauwkeurigheid.

Opgave 4. Schrijf een functie die als argumenten een vector van singuliere waarden, een rechterlid c en twee willekeurige orthogonale matrices U en V neemt. De resultaten zijn A, b, x en de theoretische waarde voor $\kappa_{b\to x}$.

Enkele Matlab commando's die van pas kunnen komen: `help`, `diag`, `rand`, `randn`, `orth`.

Opgave 5. Gebruik de functie die je gemaakt hebt om achtereenvolgens problemen op te stellen met $\kappa(A) = 1, 10^3, 10^6$. Maak gebruik van het feit dat het conditiegetal van een matrix het quotiënt van de grootste en de kleinste singuliere waarde is. Bereken voor elk probleem de relatieve perturbaties $\frac{\|\delta x\|}{\|x\|}$, $\frac{\|\delta b\|}{\|b\|}$ en de experimentele waarde $\kappa_{\exp} = \frac{\|\delta x\|}{\|x\|} / \frac{\|\delta b\|}{\|b\|}$. Beschouw telkens drie verschillende perturbaties van b:

1. Een willekeurige b en δb.
2. De vector b gelijk aan de linker singuliere vector horend bij de *grootste* singuliere waarde en de perturbatie δb gelijk aan de linker singuliere vector horend bij de *kleinste* singuliere waarde.
3. De vector b gelijk aan de linker singuliere vector horend bij de *kleinste* singuliere waarde en de perturbatie δb gelijk aan de linker singuliere vector horend bij de *grootste* singuliere waarde.

Je kan een probleem opstellen met de functie uit de vorige opgave, het rechterlid perturberen en dan het stelsel oplossen met Matlab. Een andere mogelijkheid is om zowel het oorspronkelijke als het geperturbeerde probleem op te stellen met je functie.

Je schrijft best een Matlab bestand dat je dan kan veranderen wanneer je een nieuwe Σ, b of δb wil gebruiken. Je kan ook een Matlab functie schrijven die gegeven een vector van singuliere waarden de overeenkomstige drie rijen uit de tabel hieronder genereert.

$\kappa(A)$	η	$\kappa_{b\to x}$	$\frac{\lVert\delta x\rVert}{\lVert x\rVert}$	$\frac{\lVert\delta b\rVert}{\lVert b\rVert}$	κ_{exp}
1					
10^3					
10^6					

Opgave 6. Kies δb ook eens respectievelijk loodrecht en evenwijdig op de kolomruimte van A. Wat zie je dan? Waarom is dit zo?

2.2 Conditiegetal van de coëfficiëntenmatrix

Opgave 7. Uit de formules voor de conditie van kleinste kwadratenbenaderingen blijkt dat het belangrijk is om slecht geconditioneerde coëficiëntenmatrices te vermijden. Dit is iets waar jij controle over hebt! Stel: je wil een functie benaderen door een eerstegraadsveelterm. Je beschikt over punten met als abscissen

t_1	t_2	t_3	t_4	t_5	t_6	t_7
1.01	1.02	1.03	1.04	1.05	1.06	1.07

De voor de hand liggende keuze van basis voor de lineaire veelterm is $\phi_1(t) = 1$ en $\phi_2(t) = t$. Vergelijk het gedrag van deze basis met een meer zorgvuldig geconstrueerde basis $\tilde{\phi_1}(t) = 1$ en $\tilde{\phi_2}(t) = 30(t - 1.04)$. Deze twee basissen geven aanleiding tot coëfficiëntenmatrices

$$A = \begin{bmatrix} \rule{1cm}{0.4pt} & \rule{1cm}{0.4pt} \\ \rule{1cm}{0.4pt} & \rule{1cm}{0.4pt} \\ \rule{1cm}{0.4pt} & \rule{1cm}{0.4pt} \\ \rule{1cm}{0.4pt} & \rule{1cm}{0.4pt} \\ \rule{1cm}{0.4pt} & \rule{1cm}{0.4pt} \\ \rule{1cm}{0.4pt} & \rule{1cm}{0.4pt} \\ \rule{1cm}{0.4pt} & \rule{1cm}{0.4pt} \end{bmatrix} \quad \tilde{A} = \begin{bmatrix} \rule{1cm}{0.4pt} & \rule{1cm}{0.4pt} \\ \rule{1cm}{0.4pt} & \rule{1cm}{0.4pt} \\ \rule{1cm}{0.4pt} & \rule{1cm}{0.4pt} \\ \rule{1cm}{0.4pt} & \rule{1cm}{0.4pt} \\ \rule{1cm}{0.4pt} & \rule{1cm}{0.4pt} \\ \rule{1cm}{0.4pt} & \rule{1cm}{0.4pt} \\ \rule{1cm}{0.4pt} & \rule{1cm}{0.4pt} \end{bmatrix}$$

Wat zijn de conditiegetallen? Waarom is de ene matrix beter dan de andere?