

**Numerieke Modelling en
Benadering**
Practicum 1: Eigenwaardenproblemen

De Wolf Peter
Vekemans Wout

23 april 2014

Lijst van figuren

1	Structuur van de Hessenberg matrix	4
2	Vergelijking convergentie van unshifted QR en gelijktijdig iteratie	8

Inleiding

In dit practicum onderzoeken we methoden voor het bepalen van eigenwaarden van grote matrices. In een eerste sectie beschouwen we enkele theoretische eigenschappen van de methoden. In een tweede sectie worden de convergentie-eigenschappen van de methoden onderzocht aan de hand van MATLAB-experimenten. In de derde en laatste sectie gaan we dieper in op één van de methoden, namelijk de Jacobi-methode.

Theoretische eigenschappen

Opgave 1

Gelijktijdige iteratie is een andere naam voor het toepassen van de methode van de machten op meerdere kolommen van A tegelijk. Zo kunnen de grootste n eigenwaarden van de matrix A gevonden worden. Om met gelijktijdige iteratie de kleinste n eigenwaarden van A te berekenen is het voldoende om het algoritme toe te passen op A^{-1} in plaats van op A . Dit is makkelijk in te zien. De eigenwaarden van de inverse van een singuliere matrix A zijn gelijk aan het omgekeerde van de eigenwaarden van A .

$$Ax = \lambda x \quad (1a)$$

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

$$A^{-1} A \frac{1}{\lambda} x = A^{-1} x \quad (1c)$$

$$A^{-1} x = \frac{1}{\lambda} x \quad (1d)$$

Als we dit algoritme uitschrijven krijgen we volgende pseudocode:

```
Initialiseer  $\hat{Q}^{(0)} \in \mathbb{R}^{m \times n}$  met  $n$  orthonormale kolommen.  
for  $k = 1, 2, \dots$  do  
   $Z = A^{-1} \hat{Q}^{k-1}$   
   $\hat{Q}^k \hat{R}^k = Z$   
end for
```

Dit zal convergeren naar de n grootste eigenwaarden van A^{-1} , die zoals eerder aangetoond in (1) overeenkomen met de n kleinste eigenwaarden van A .

Opgave 2

a) De Rayleigh quotiënt iteratie convergeert zeer snel naar de juiste waarde van de eigenwaarden van A . Elke iteratiestap verdrievoudigt het aantal juiste cijfers. Elke stap dient er wel een stelsel te worden opgelost dat meer en meer singulier wordt naarmate de benadering voor de eigenwaarde juist wordt. Het betreft het stelsel

$$(A - \lambda^{(k-1)}I)w = v^{(k-1)} \quad (2)$$

Als λ dichterbij een eigenwaarde van A komt wordt $A - \lambda^{(k-1)}I$ meer en meer singulier, en zal de inverse van deze matrix, die nodig is om (2) op te lossen, ervoor zorgen dat de vector w enorm wordt opgeblazen. De vector zal echter wel in de juiste richting blijven wijzen, waardoor dit geen problemen oplevert.

b) Gegeven een matrix $A \in \mathbb{R}^{m \times m}$ en een vector $x \in \mathbb{R}^{m \times 1}$, waarbij x een benadering is voor een eigenvector van A . Toon aan dat de oplossing $\rho \in \mathbb{R}$ van het minimalisatieprobleem

$$\min_{\rho \in \mathbb{R}} \|Ax - \rho x\|_2 \quad (3)$$

overeenkomt met het Rayleighquotient van x .

Dit probleem lijkt op het standaard kleinste kwadraten probleem $Ax \approx b$. We proberen nu een ρ te vinden zodat $\|Ax - \rho x\|_2$ geminimaliseerd wordt. De overeenkomstige normaalvergelijkingen (cf $A^T Ax = A^T b$) worden gegeven door

$$x^T x \rho = x^T Ax \quad (4a)$$

$$\rho = \frac{x^T Ax}{x^T x} = r(x) \quad (4b)$$

waarbij $r(x)$ het Rayleigh quotient van x voorstelt.

Opgave 3

Convergentie-experimenten

Voor al deze experimenten wordt gebruik gemaakt van een volle, reële symmetrische matrix A . We laden hiervoor de gegeven matrix *mat1.txt* in in MATLAB.

Opgave 4

Het uitvoeren van het *spy*-commando laat zien dat er geen enkel *non-zero* element in de matrix zit. De uitvoering van het QR-algoritme op een matrix met die afmetingen (nog steeds relatief klein) zou zeer veel werk vragen.

Na het reduceren tot Hessenberg vorm zien we dat alle elementen onder de eerste benedendiagonaal nul zijn geworden. Verdere analyse leert ons dat ook de elementen boven de eerste bovendiagonaal allemaal van grootteorde ϵ_{mach} zijn. Als we de Hessenberg vorm van de matrix afronden tot op 15 decimalen verkrijgen we een tridiagonale matrix. (zie figuur 1) Dit komt doordat de originele matrix symmetrisch was. De reductie naar Hessenberg vorm zorgt ervoor dat het algoritme niet op een volledige matrix moet inwerken. Het uitvoeren van het QR-algoritme vraagt nu slechts $\mathcal{O}(n^2)$ flops.



Figuur 1: Structuur van de Hessenberg matrix

Opgave 5

Het toepassen van de drie verschillende versies van het QR algoritme geeft ons figuur ???. We zien duidelijk dat er 1 methode trager is dan de rest. Als maat voor de convergentie gebruiken we de relatieve fout op de berekende eigenwaarden.

Voor het algoritme zonder shifts stemt dit volledig overeen met de theorie. Theorem 28.4 (p. 218 in het handboek) geeft duidelijk aan dat bij het gebruik van QR zonder shifts $A^{(k)}$ lineair convergeert naar een diagonaalmatrix met de eigenwaarden van A als elementen. In Tabel 1 staan de relatieve fouten op de vijf dominante eigenwaarden van de matrix, bij gebruik van het algoritme zonder shifts. Zie ook figuur ?? voor een grafische illustratie van de convergentie.

Voor Wilkinson shift zien we in tabel 3 dat er over het algemeen snellere convergentie is dan bij het algoritme zonder shifts. De verwachte kubische convergentie blijkt niet echt duidelijk uit de resultaten.

Iteratie	23	22	21	20	19
5	3.4173E-02	7.7106E-02	7.8187E-02	6.4348E-02	1.6090E-02
10	1.7414E-02	1.5644E-02	5.4016E-02	2.7533E-02	3.4097E-03
15	1.1828E-02	7.6086E-03	2.5512E-02	2.6705E-02	1.8666E-02
20	8.1546E-03	7.6596E-03	1.4755E-02	1.9487E-02	2.2061E-02
25	5.5253E-03	5.5328E-03	9.3537E-03	1.1783E-02	1.6347E-02
30	3.6818E-03	3.7494E-03	5.9809E-03	7.2101E-03	1.1119E-02
35	2.4229E-03	2.4821E-03	3.7863E-03	4.3930E-03	7.2111E-03
40	1.5805E-03	1.6235E-03	2.3726E-03	2.6464E-03	4.5407E-03
45	1.0249E-03	1.0543E-03	1.4757E-03	1.5767E-03	2.8068E-03
50	6.6195E-04	6.8157E-04	9.1340E-04	9.3132E-04	1.7155E-03
55	4.2643E-04	4.3936E-04	5.6357E-04	5.4674E-04	1.0413E-03
60	2.7425E-04	2.8270E-04	3.4703E-04	3.1955E-04	6.2951E-04
65	1.7618E-04	1.8169E-04	2.1342E-04	1.8615E-04	3.7964E-04
70	1.1310E-04	1.1668E-04	1.3115E-04	1.0814E-04	2.2861E-04
75	7.2573E-05	7.4891E-05	8.0558E-05	6.2672E-05	1.3754E-04
80	4.6554E-05	4.8055E-05	4.9466E-05	3.6236E-05	8.2704E-05
85	2.9857E-05	3.0829E-05	3.0368E-05	2.0900E-05	4.9712E-05
90	1.9147E-05	1.9775E-05	1.8641E-05	1.2024E-05	2.9873E-05
95	1.2277E-05	1.2683E-05	1.1442E-05	6.8983E-06	1.7948E-05

Tabel 1: Relatieve fout op de vijf grootste berekende eigenwaarden (zonder shifts)

Iteratie	23	22	21	20	19
5	2.7954E-02	1.1750E-01	1.0582E-01	2.4273E-01	2.3798E-01
10	1.4943E-02	2.0071E-02	8.4508E-02	8.2623E-02	1.0694E-01
15	9.6654E-03	3.1292E-03	9.0310E-02	4.8149E-02	2.6853E-02
20	6.1540E-03	3.7826E-03	3.4726E-02	4.6832E-02	2.3820E-02
25	3.8228E-03	2.6950E-03	1.3951E-02	4.8232E-02	1.2391E-02
30	2.3320E-03	1.7462E-03	7.0906E-03	1.6654E-02	3.8615E-02
35	1.4057E-03	1.0889E-03	3.9605E-03	3.2062E-03	4.6617E-02
40	8.4102E-04	6.6583E-04	2.2457E-03	1.7235E-04	3.9414E-02
45	5.0085E-04	4.0271E-04	1.2673E-03	6.4671E-04	2.6736E-02
50	2.9744E-04	2.4205E-04	7.1057E-04	5.1331E-04	2.6268E-02
55	1.7635E-04	1.4495E-04	3.9652E-04	3.2931E-04	2.3048E-02
60	1.0445E-04	8.6613E-05	2.2060E-04	1.9614E-04	1.4949E-02
65	6.1824E-05	5.1682E-05	1.2251E-04	1.1316E-04	8.8874E-03
70	3.6582E-05	3.0811E-05	6.7957E-05	6.4321E-05	4.9702E-03
75	2.1642E-05	1.8357E-05	3.7671E-05	3.6294E-05	2.6765E-03
80	1.2801E-05	1.0932E-05	2.0873E-05	2.0403E-05	1.4106E-03
85	7.5716E-06	6.5079E-06	1.1561E-05	1.1446E-05	7.3479E-04
90	4.4782E-06	3.8731E-06	6.4022E-06	6.4144E-06	3.8037E-04
95	2.6485E-06	2.3044E-06	3.5445E-06	3.5920E-06	1.9627E-04

Tabel 2: Relatieve fout op vijf grootste berekende eigenwaarden (Rayleigh shift)

Iteratie	23	22	21	20	19
5	3.5241E-02	1.7599E-01	1.5860E-01	2.9753E-01	2.7171E-01
10	1.6399E-02	3.5473E-02	1.0074E-01	7.4094E-02	1.4469E-01
15	1.0546E-02	1.9560E-03	8.6089E-02	5.8539E-02	3.5689E-02
20	6.7499E-03	3.9381E-03	4.2712E-02	4.9227E-02	2.3114E-02
25	4.2117E-03	2.9159E-03	1.6377E-02	5.5393E-02	6.6471E-03
30	2.5772E-03	1.9121E-03	8.0228E-03	2.1555E-02	3.4651E-02
35	1.5566E-03	1.1991E-03	4.4386E-03	4.7371E-03	4.2576E-02
40	9.3243E-04	7.3546E-04	2.5159E-03	1.3738E-04	4.3532E-02
45	5.5571E-04	4.4559E-04	1.4217E-03	6.3519E-04	2.8144E-02
50	3.3017E-04	2.6809E-04	7.9809E-04	5.5136E-04	2.5672E-02
55	1.9580E-04	1.6064E-04	4.4572E-04	3.6282E-04	2.4825E-02
60	1.1599E-04	9.6021E-05	2.4810E-04	2.1829E-04	1.6428E-02
65	6.8662E-05	5.7309E-05	1.3782E-04	1.2651E-04	9.9183E-03
70	4.0631E-05	3.4171E-05	7.6463E-05	7.2069E-05	5.6034E-03
75	2.4038E-05	2.0361E-05	4.2390E-05	4.0710E-05	3.0357E-03
80	1.4219E-05	1.2126E-05	2.3489E-05	2.2898E-05	1.6053E-03
85	8.4102E-06	7.2194E-06	1.3012E-05	1.2851E-05	8.3769E-04
90	4.9742E-06	4.2968E-06	7.2056E-06	7.2026E-06	4.3405E-04
95	2.9419E-06	2.5566E-06	3.9895E-06	4.0338E-06	2.2408E-04

Tabel 3: Relatieve fout op vijf grootste berekende eigenwaarden (Wilkinson)

Opgave 6

De Rayleigh quotiënt iteratie en de Rayleigh shift methode zouden allebei kubisch moeten convergeren. In figuur ?? zijn onze numerieke bevindingen weergegeven. We zien duidelijk dat er een gelijkaardige convergentiesnelheid is. Dit komt doordat de Rayleigh shift methode gebruik maakt van het Rayleigh quotiënt. Het QR algoritme met Rayleigh shifts is namelijk een methode waarbij er snelle convergentie is in de laatste kolom van $Q^{(k)}$. Het Rayleigh quotiënt $\mu^{(k)}$ van de laatste kolom van de matrix is dan een goede schatting voor de eigenwaarde horende bij de laatste kolom.

$$\mu^{(k)} = \frac{(q_m^{(k)})^T A q_m^{(k)}}{(q_m^{(k)})^T q_m^{(k)}} = (q_m^{(k)})^T A q_m^{(k)} \quad (5)$$

Nu is dit getal gelijk aan $A_{mm}^{(k)}$ en kan dit getal zo als shift gebruikt worden voor de volgende iteratiestap.

De convergentie van gelijktijdige iteratie en QR zonder shifts wordt geïllustreerd in figuur 2. Als residu gebruiken we hier de relatieve fout op de diagonaal van A . Dit kan een licht vertekend beeld geven als er enkele eigenwaarden niet zo snel convergeren, maar voor een globaal beeld is het voldoende. Het is duidelijk dat deze methodes ongeveer even snel convergeren. Dit kan eenvoudig worden aangetoond. Het algoritme voor gelijktijdige iteratie is als volgt:

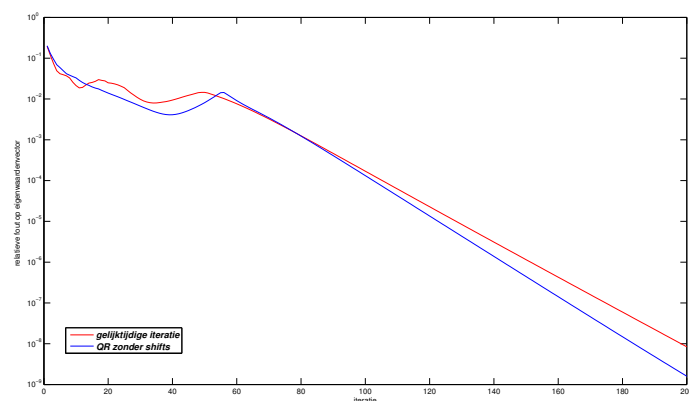
```

Kies  $\hat{Q}^{(0)} \in \mathbb{R}^{m \times n}$  met orthonormale kolommen
for  $k = 1, 2, \dots$  do
     $Z = A\hat{Q}^{(k-1)}$ 
     $\hat{Q}^{(k)}\hat{R}^{(k)} = Z$ 
end for

```

Theorema 28.3 (p. 216) in het handboek toont aan dat gelijktijdige iteratie en unshifted QR dezelfde matrices $\underline{R}^{(k)}, \underline{Q}^{(k)}$ en $\underline{A}^{(k)}$ genereren. Het is ook eenvoudig in te zien dat dit de matrices zijn die gedefinieerd worden door de k^{de} macht van de QR factorisatie van A .

Volgens theorema 28.4 (p. 218) convergeert het QR algoritme zonder shifts lineair naar een diagonaalmatrix. Aangezien gelijktijdige iteratie dezelfde matrices produceert moet deze methode een gelijkaardige convergentiesnelheid hebben. Uit figuur 2 wordt duidelijk dat matrix A inderdaar lineair convergeert naar een diagonaalmatrix.



Figuur 2: Vergelijking convergentie van unshifted QR en gelijktijdig iteratie

Opgave 7

Alternatieve eigenwaardenalgoritmen

Opgave 8

Opgave 9

Opgave 10