

Iteratieve methoden

 tuyaux.winak.be/index.php/iteratieve_methoden

Calculus

Richting Wiskunde

Jaar 2BWIS

Inhoud van het vak en examen

Het vak iteratieve methodes (voor het oplossen van lineaire stelsels en eigenwaarden problemen) tracht het probleem $Ax=b$ op te lossen zonder de inverse van A te berekenen. Dit is nuttig voor zeer grote stelsels, omdat de kost om $A^{-1}A^{-1}$ te berekenen stijgt met ongeveer $O(n^3)$. Tenslotte worden er enkele methodes beschouwt om eigenwaarden van matrices iteratief te berekenen.

De cursus bestaat uit 2 types van methoden: de stationaire methoden en de Krylov methoden. Voor dit vak wisselen de theorie en de oefeningen elkaar ongeveer wekelijks af. De oefeningen zijn vooral bedoeld om de theorie te implementeren en inzicht te krijgen in de algoritmes.

Doorheen het jaar zijn er enkele taken waar je een oefening moet oplossen, vaak in Matlab, en een verslag moet inleveren. Deze taken worden gequoteerd. Het examen bestaat enkel uit theorie, waarbij veel belang wordt gehecht aan verschillen en overeenkomsten tussen verschillende methoden.

Examenvragen

Juni 2023

1. (5p) Gegeven een convergerende stationaire iteratie
 $x(k+1) = Mx(k) + c$ met $\rho(M)$ de spectrale straal en M een ijle matrix met maximaal K niet-nul elementen per rij.
 - a) Wat is het verwachte aantal iteraties om de relatieve fout $\|e(k)\|_2 / \|e(0)\|_2$ onder een tolerantie τ te krijgen?
 - b) Geef een schatting van het aantal elementaire operaties (vermenigvuldigingen of optellingen van twee getallen) die nodig zijn om deze tolerantie te bereiken.
 - c) Wat is het concept achter de Chebychev-versnelling van deze iteratie? Waarom zou dit sneller convergeren? Kan je een schatting geven van de verwachte convergentiesnelheid? Hoeveel iteraties verwacht je om de (relatieve) fout onder de tolerantie τ te krijgen?
 - d) Zal een Krylov methode beter dan een stationaire iteratie of Chebychev-versnelling convergeren? Motiveer je antwoord.

2. (5p) Gegeven een symmetrische en positief definitie matrix $A \in \mathbb{R}^{n \times n}$.
 Definieer de conjugate gradient (CG) methode als

$$x_k = \arg\min_{x \in K_k(A, r_0)} \|Ax - b\|_A$$
 waar $K_k(A, r_0)$ een Krylov ruimte voorstelt.
 - a) Waarom staan de opeenvolgende residuën loodrecht?
 - b) Wat is de recursierelatie voor de residuën? Waarom gebruiken we deze recursierelatie in plaats van een expliciete berekening?
 - c) Waarom vinden we, als we exacte berekeningen aannemen, een iteratie in maximaal n iteraties?
 - d) Wat gebruikt een CG-methode dat een stationaire en Chebychev-methode niet gebruiken?

3. (5p) Stel $A \in \mathbb{R}^{n \times n}$ een niet-singuliere matrix.
 - a) In GMRES gebruiken we de Arnoldi relatie. Wat is deze relatie? Wat is een Hessenberg matrix en waarom komt deze speciale structuur voor in de Arnoldi relatie? Toon aan. Hoe kunnen we de elementen van deze Hessenberg matrix uitrekenen?
 - b) Wat is een happy breakdown en waarom hebben we dan een oplossing gevonden? Wat is het verschil met een serious breakdown die in BiCG voorkomt? Wanneer komt zo een serious breakdown voor?

4. (5p) Stel $A \in \mathbb{R}^{n \times n}$ een niet-vierkante matrix. De Golub-Kahan iteratie veralgemeent iteratieve methoden naar niet-vierkante matrices. Ze wordt onder meer gebruikt om de singular value decomposition (SVD) van een matrix uit te rekenen en de normale vergelijking $ATAx=ATb$ op te lossen. We kunnen de volgende relatie opstellen

$$AV_k = U_k B_k e_k,$$

$$AV_k = U_k B_k e_k,$$

waarbij $V_k^T V_k = I_k$ en $U_k^T U_k = I_k$ twee orthogonale basissen zijn.

a) Wat kan je vertellen over de structuur van B_k ? Waar zijn de nul elementen?

Als je veronderstelt dat B_k een onderdriehoeksmatrix is, wat is dan de structuur van B_k ? Kan je de matrix elementen van B_k uitrekenen?

b) Kan je een recursierelatie opstellen voor de vectoren v_k , de kolommen in V , en u_k , de kolommen in U , en de matrixelementen van B_k ?

Juni 2021

1. Waarom convergeert CG voor $A \in \mathbb{R}^{n \times n}$ symmetrisch en positief definit, in maximaal n iteraties? Wat zijn Ritz-waarden van A ? Hoe kan je uitrekenen? Waarom is een Ritz-waarde een nulpunt van de minimale veelterm van CG? Toon aan. (5 punten)
2. Vergelijk GMRES met BiCG voor een niet-symmetrische matrix $A \in \mathbb{R}^{n \times n}$. (6 punten)
 1. Maak een schets van het typisch verloop van de 2-norm van de residuen voor GMRES en BiCG als functie van het aantal iteraties. Wat is de relatieve positie van de convergerende curve voor GMRES tegenover BiCG? Welke curven zijn strikt dalend als functie van het aantal iteraties?
 2. Op welk minimalisatieprincipe is elk algoritme gebaseerd?
 3. Op de volgende pagina's staan beide algoritmes. Bespreek de verschillen tussen beide algoritmes. Wat zijn de belangrijkste bouwstenen? Welke matrixbewerkingen, recursies en producten komen voor? Wat zijn de verschillen?
 4. Wat zijn de mogelijke breakdowns in beide algoritmes en wat betekenen ze voor de convergentie?
 5. Bespreek de voor- en nadelen voor elk algoritme. Wanneer gebruik je GMRES of BiCG?

3. Chebyshev vs. Stationaire methode (4 punten)

1. Stel $x(k+1) = Mx(k) + c$ met $M \in \mathbb{R}^{n \times n}$ een stationaire iteratie. Als je deze stationaire iteratie toepast op een initiële fout die een eigenvector is van M , hoe verandert de fout dan?
 2. Hoe evolueert de fout wanneer je de Chebyshev iteratie toepast op een eigenvector van M ?
 3. Stel dat de spectrale straal van M gelijk is aan $\rho < 1$. Stel een willekeurige initiële fout. Maak een tekening van de evolutie van de fout voor zowel stationaire iteratie als Chebyshev iteratie. Wat zijn de hellingen van de convergente curven van beide methoden?
4. Waarom staan de zoekrichtingen A-loodrecht in CG? Toon aan. (5 punten)

Januari 2018

1. Wanneer gebruik je BiCG? Kan je het idee uitleggen achter de afleiding van BiCG? Schrijf niet heel de afleiding van BiCG op, enkel het belangrijkste idee. Wat zijn de voordelen, wat zijn de nadelen van BiCG? Wat is een serious breakdown? Kan je wiskundig uitleggen wanneer en waarom dit voorkomt?
2. Geef een voorbeeld van een stationaire methode en geef een voldoende voorwaarde opdat de methode convergeert voor een willekeurige initiële schatting.
3. Stel A symmetrisch en positief definit, en b een willekeurige vector. Vergelijk de performantie bij het oplossen van $Ax = b$ van CG en Chebyshev in 2 grafieken. Eén grafiek vergelijkt de fout in de A -norm van de 2 methoden en één grafiek vergelijkt de opeenvolgende residuen van de 2 methoden. Let op de relatieve positie van de 2 curven van Chebyshev en CG. Motiveer je keuzes.
4. Hoe kies je de parameters van Chebyshev iteraties? Waar houd je rekening mee?
5. Wat is een happy breakdown bij GMRES? Wanneer komt die voor? Waarom noemt men dit een happy breakdown?
6. Waar/Vals + waarom:
 - Een Hessenbergmatrix van een symmetrische matrix is ook symmetrisch.
 - De eigenwaarden van de Hessenbergmatrix na N iteraties van GMRES komen overeen met de eigenwaarden van A , een $N \times N$ matrix.
 - Door GMRES te herstarten profiteer je van superlineaire convergentie.
 - Eigenwaardenberekeningen kan je versnellen door preconditioning.

Augustus 2017

1. Bespreek de verschillen in gebruik, werking en convergentie tussen CG, GMRES en BiCGstab. Wees beknopt. Als je de convergentie van deze drie methoden op stelsel met een symmetrische en positief definitie matrix zou vergelijken, hoe zouden de plots voor residu en fout eruit zien?
2. Wat weet je over de zoekrichtingen in CG?
3. Waarom kan je in iteratieve methoden het residu na k stappen schrijven als polynoom van de matrix toegepast op een vector? Wat zijn de eigenschappen van de polynoom? Hoe kan je de polynoom gebruiken om convergentie te analyseren? Hoe ziet deze polynoom eruit voor CG?

Juni 2016

1. In CG wordt het residu niet expliciet uitgerekend. Er wordt eerder een recursie gebruikt. Dit vermijdt het uitrekenen van een extra matrix-vector product.
 - Wat is de recursie voor het residu? Hoe vind je deze?
 - Omdat een computer rekent met een eindige precisie zijn er afrondingsfouten bij het berekenen van deze recursie. Stel een recursie op voor de afwijking tussen het residu uit het algoritme en het echte residu.
 - Kan je de groei van deze fout schatten?
2. Stel A een asymmetrische, niet singuliere matrix. De matrix heeft complexe eigenwaarden met de volgende eigenschappen. Er liggen 5 eigenwaarden in cirkel met straal 1 en middelpunt $-1.5 + 0i$ en 1000 eigenwaarden in een cirkel met straal 40 en middelpunt $50 + 0i$. We lossen het stelsel $Ax=b$ op, waarbij b een lineaire combinatie is van alle eigenvectoren. We gebruiken GMRES, BiCG en CGS.
 - Maak een goede kwantitatieve voorspelling van het verwachte convergentiegedrag voor de drie methoden in functie van het aantal iteraties en het aantal matrix-vector producten.
 - Maak een schets van het convergentiegedrag
 - Geef ook een schatting van de relatieve rekentijd. Welke methode is het snelste? Argumenteer.
 - Als er nu 10 eigenwaarden zijn in plaats van 5 in de eerste cirkel en 2000 in de tweede cirkel, wat verandert er dan aan het convergentiegedrag?
3. Wat is een happy breakdown bij GMRES en een breakdown bij BiCG? Wanneer komen ze voor?
4. Wat is een Ritz-paar? Waarvoor kan je het gebruiken?
5. Voor wat kan je een Rayleigh Quotiënt gebruiken?
6. Wanneer zou je een directe methode verkiezen? Wanneer een iteratieve methode?

Augustus 2016

1. Wat is een Hessenberg matrix en in welk algoritme komt zo'n matrix voor?
2. Waarom heeft zo een Hessenberg matrix die speciale structuur? (Toon dit aan via de Arnoldi relaties)
3. Geef een voorbeeld matrix waarbij je verwacht dat Toegevoegde Gradiënten superlineaire convergentie zal vertonen. Leg uit waarom deze matrix superlineaire convergentie zal vertonen. Komt superlineaire convergentie ook voor bij Chebyshev iteraties? Motiveer!
4. Wanneer kan je Fast Fourier Preconditioning toepassen? Van welke eigenschap van Fourier transformaties maken we gebruik?
5. Waarom vertonen bi-orthogonale iteratieve methoden soms een onregelmatig convergentiegedrag? Hoe kan je deze onregelmatigheden verzachten?
6. Wat is een benaderende inverse? Wat is het verband met preconditioning?

2015

1. Bespreek de verschillen in gebruik, convergentie en werking van GMRES en CG.
Antwoord beknopt.
2. Waarom kunnen we in CG de schattingen met elkaar in verband brengen door de vorige schatting en een bepaalde zoekrichting? Hoe bepalen we deze zoekrichting?
3. Hoe kunnen we bij iteratieve methoden het residu na k stappen in verband brengen met een polynoom van de matrix en het initiële residu? Welke polynoom is dit voor CG?
4. Geef drie voorbeelden van preconditioners. Leg kort uit wat ze doen.
5. We bekijken het 2D Poisson probleem en passen opnieuw eindige differenties toe. Op deze matrix gebruiken we CG en de Richardson iteratie. Welke convergeert het best? Als we het aantal roosterpunten verdubbelen langs beide zijden, wat is dan het resultaat voor de convergentie?

2012

1. Stel A een $N \times N$ diagonalseerbare matrix met eigenwaarden $\lambda_j = 10 + 9.2 \exp(i2\pi j/N)$ met $j=1, \dots, N$. Figuur 1 (figuur gegeven op examen) toont de convergentie van $\|r_k\|_2$ voor GMRES en GMRES(10) toegepast op de matrix A met $N=300$. De rechterhand is een willekeurige vector en de startvector is de nulvector.
 - Welke curve uit de figuur is GMRES en welke curve is GMRES(10)?
 - Geef een scherpe bovengrens voor de convergentie voor elke curve? Motiveer je antwoord.
 - Als je de convergenties zou kunnen verderzetten voor een groter aantal iteraties, hoe zouden ze eruit zien?
 - Is er superlineaire convergentie?
2. Stel $\min \|Ax - b\|_2$ een kleinste kwadraten probleem met $A \in \mathbb{R}^{m \times n}$ een $m \times n$ rechthoekige matrix met $m \geq n$ en de matrix heeft volledige kolom rang. Dit probleem kan opgelost worden door CG toe te passen op de normaalvergelijking $A^T A x = A^T b$. Deze methode wordt CGLS genoemd naar Conjugate Gradients Least-Squares.
 - Stel x_k de oplossing van bovenstaand probleem en r_k het corresponderend residu. Toon aan dat CGLS inderdaad $\|r_k\|_2$ minimaliseert.
 - Wat kan je zeggen over opeenvolgende $\|r_k\|_2$?
 - Wat is de relatie tussen de zoekrichtingen in CGLS en in CG?
3.
 - Bespreek de voor- en nadelen van BiCG?
 - Wat is BiCGSTAB? Zijn er voordelen ten opzichte van BiCG?

Categorieën:

- Wiskunde
- 2BWIS