# 10. Jacobi-Davidsonova metoda

Iščemo eno ali več (dominantnih) lastnih vrednosti $n \times n$ matrike $A$, kjer je $n$ velik. Matrika $A$ je lahko razpršena, ni pa nujno.

Zaradi velikega $n$ reševanje sistema z matriko $A$ ne pride v poštev.

Pri metodah, ki temeljijo na podprostorih Krilova (Arnoldi, Lanczos, ...) je podprostor, v katerem iščemo aproksimacije (Ritzeve, Petrove, harmonične Ritzeve vrednosti, ...) v bistvu odvisen le od začetnega vektorja $v_1$.

Pri Jacobi-Davidsonovi metodi podprostor širimo na drug način. Metodo sta razvila van der Vorst in Sleijpen leta 1996. Temelji pa na kombinaciji dveh idej:

- Jacobijeva metoda (1840)
- Davidsonova metoda (1975)

# Davidsonova metoda

Naj vektorji $v_1, \ldots, v_k$ tvorijo ortonormirano bazo podprostora $\mathcal{V}_k$ dimenzije $k$ in naj bo $\theta_k$ največja Ritzeva vrednost za $A$ in $\mathcal{V}_k$ z ustreznim Ritzevim vektorjem $u_k$.

V kateri smeri naj razširimo podprostor $\mathcal{V}_k$ na $\mathcal{V}_{k+1}$, da bomo dobili še boljše približke za dominantno lastno vrednost?

Naj bo

$$r = Au_k - \theta_k u_k.$$

Davidsonova ideja je, da rešimo sistem

$$(D_A - \theta_k I)t = r,$$

kjer je $D_A$ diagonala matrike $A$. Z vektorjem $t$ (ki ga ortonormiramo glede na prejšnje vektorje) razširimo $\mathcal{V}_k$.

Metoda v praksi deluje zelo dobro za iskanje dominantnih lastnih vrednosti diagonalno dominantnih matrik. Matematični temelji pa niso tako solidni!

Matrike $D_A - \theta_k I$ si ne smemo predstavljati kot aproksimacije za $A - \theta_k I$. Če namreč namesto $D_A - \theta_k I$ vzamemo $A - \theta_k I$, dobimo $t = u_k$, ker pa je $u_k \in \mathcal{V}_k$, s tem ne moremo razširiti baze.

# Algoritem za Davidsonovo metodo

izberi vektor $\|v_1\| = 1$
$k = 1, \ldots, m$
  izračunaj $k$-to vrstico in $k$-ti stolpec $B_k = V_k^H A V_k$.
  izračunaj največjo Ritzevo vrednost $\theta_k$ in pripadajoči Ritzev vektor $u_k = V_k s_k$
  $r = Au_k - \theta_k u_k$
  $t = (D_A - \theta_k I)^{-1} r$
  vektor $t$ ortogonaliziraj glede na $v_1, \ldots, v_k$ in ostanek vzemi za $v_{k+1}$

Namesto $(D_A - \theta_k I)^{-1}$ bi lahko uporabili katero koli drugo aproksimacijo za $(A - \theta_k I)^{-1}$.

Za ortogonalizacijo vektorja $t$ uporabimo MGS oziroma še bolje RGS.

Če za aproksimacijo $(A - \theta_k I)^{-1}$ vzamemo kar $(I - \theta_k I)^{-1}$, dobimo Arnoldijevo metodo. To očitno dobimo tudi, kadar ima $A$ konstantno diagonalo.

Matrika $B_k$ za razliko od Arnoldijeve metode nima nobene posebne oblike in je polna.

Za razliko od Arnoldijeve metode tukaj računamo le eno lastno vrednost.

# Jacobijeva metoda ortogonalnih popravkov

Naj bo $a_{11} = \alpha$ največji diagonalni element diagonalno dominantne matrike $A$. Potem je $\alpha$ približek za dominantno lastno vrednost $\lambda$, približek za lastni vektor pa je $e_1$.

Lastni problem

$$A \begin{bmatrix} 1 \\ z \end{bmatrix} = \begin{bmatrix} \alpha & c^T \\ b & F \end{bmatrix} \begin{bmatrix} 1 \\ z \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ z \end{bmatrix},$$

kjer je $F$ kvadratna matrika, $\alpha$ je skalar, $b, c, z$ pa vektorji reda $n-1$, je ekvivalenten

$$\begin{aligned} \lambda &= \alpha + c^T z, \\ (F - \lambda I)z &= -b \end{aligned}$$

Jacobi je predlagal iterativno reševanje, kjer vzamemo $z_1 = 0$ in

$$\begin{aligned} \theta_k &= \alpha + c^T z_k, \\ (D - \theta_k I)z_{k+1} &= (D - F)z_k - b, \end{aligned}$$

kjer je $D$ diagonala $F$.

V vseh korakih Jacobijevega pristopa iščemo lastni vektor oblike $u = \begin{bmatrix} 1 \\ z \end{bmatrix}$, torej iščemo ortogonalne popravke $\begin{bmatrix} 0 \\ z \end{bmatrix} = u - (e_1^H u)e_1$ za začetni približek $e_1$.

Ne upoštevamo, da med postopkom dobimo boljše približke $u_k = \begin{bmatrix} 1 \\ z_k \end{bmatrix}$ in da bi potem lahko iskali ortogonalne popravke $u - (u_k^H u)u_k$ za $u_k$.

# Podobnosti med Jacobijevo in Davidsonovo metodo

Vzemimo matriko $A = \begin{bmatrix} \alpha & c^T \\ b & F \end{bmatrix}$ kot pri Jacobiju in naj bo $v_1 = e_1$. Vektor $u_k$ skaliramo tako, da je $u_k = \left[\begin{smallmatrix} 1 \\ z_k \end{smallmatrix}\right]$.

Naj bo $\theta_k$ aproksimacija za lastno vrednost. Za ostanek velja

$$r_k = (A - \theta_k I) u_k = \begin{bmatrix} \alpha - \theta_k + c^T z_k \\ (F - \theta_k I) z_k + b \end{bmatrix}.$$

Po Davidsonu $t_k$ dobimo iz $(D_A - \theta_k I) t_k = -r_k$. Če označimo $t_k = \left[\begin{smallmatrix} * \\ y_k \end{smallmatrix}\right]$, velja

$$(D - \theta_k I) y_k = -(F - \theta_k I) z_k - b = (D - F) z_k - (D - \theta I) z_k - b,$$

oziroma

$$(D - \theta_k I)(z_k + y_k) = (D - F) z_k - b.$$

Vektor $z_k + y_k$ se tako ujema z $z_{k+1}$, ki ga dobimo po enem koraku Jacobijeve metode, če začnemo z $z_k$.

# Razlike med Jacobijevo in Davidsonovo metodo

- Jacobi: Kot naslednji približek za lastni vektor vzamemo $\widehat{u}_{k+1} = \left[ \begin{smallmatrix} 1 \\ z_k + y_k \end{smallmatrix} \right] = u_k + \widehat{y}_k$. Torej vzamemo komponento $\widehat{y}_k = \left[ \begin{smallmatrix} 0 \\ y_k \end{smallmatrix} \right]$ vektorja $t_k$, ki je ortogonalna na $u_1$ in za naslednji približek vzamemo $u_{k+1} = u_k + \widehat{y}_k$ in $\theta_{k+1} = e_1^H A u_{k+1}$.
- Davidson: Vzamemo komponento $v_{k+1}$ vektorja $t_k$, ki je ortogonalna na $u_1, \ldots, u_k$ in razširimo prostor na $\mathcal{V}_{k+1} = \mathcal{L}(v_1, \ldots, v_k, v_{k+1})$, potem pa vzamemo za $\theta_{k+1}, u_{k+1}$ Ritzev par matrike $A$ za $\mathcal{V}_{k+1}$.

Čeprav velja $\mathcal{L}(u_1, \ldots, u_k, \widehat{u}_{k+1}) = \mathcal{L}(u_1, \ldots, u_k, u_{k+1})$ se metodi ne ujemata, saj približki $\theta_k$ niso enaki, to pa pomeni, da se razlikujejo tudi popravki $t_j$.

V obeh metodah nove smeri iščemo s fiksnimi operatorji ($D_A$ pri Davidsonu oz. $D$ pri Jacobiju), ne pa s takimi, ki bi bili povezani s prostorom v katerem iščemo popravek oz. novo smer. To popravimo v Jacobi-Davidsonovi metodi, ki temelji na obeh metodah.

# Jacobi-Davidsonova metoda

Naj bo $\mathcal{V}_k$ podprostor dimenzije $k$ in naj bo $\theta_k$ dominantna Ritzeva vrednost za $A$ in $\mathcal{V}_k$ z ustreznim normiranim Ritzevim vektorjem $u_k$.

Sedaj popravek za $u_k$ iščemo v podprostoru $u_k^\perp$. Ortogonalna projekcija $A$ na ta podprostor je $B = (I - u_k u_k^H)A(I - u_k u_k^H)$. Zapišemo (upoštevamo $\theta_k = u_k^H A u_k$)

$$A = B + A u_k u_k^H + u_k u_k^H A - \theta_k u_k u_k^H.$$

Nastavek je $A(u_k + v) = \lambda(u_k + v)$, od tod pa iz $B u_k = 0$ sledi

$$(B - \lambda I)v = -r + (\lambda - \theta_k - u_k^H A v)u_k.$$

Ker sta leva stran in $r$ ortogonalna na $u_k$, mora biti faktor pred $u_k$ enak $0$ in dobimo

$$(B - \lambda I)v = (I - u_k u_k^H)(A - \lambda I)(I - u_k u_k^H)v = -r.$$

Kot pri Jacobijevi in Davidsonovi metodi namesto $\lambda$ vzamemo približek $\theta_k$.

Korekcijsko enačbo

$$(I - u_k u_k^H)(A - \theta_k I)(I - u_k u_k^H)v = -r$$

rešimo le približno in z novim vektorjem razširimo bazo $\mathcal{V}_k$.

# Jacobi-Davidsonova metoda - 2

Naj bo $\mathcal{V}_k$ podprostor dimenzije $k$ in naj bo $\theta_k$ dominantna Ritzeva vrednost za $A$ in $\mathcal{V}_k$ z ustreznim normiranim Ritzevim vektorjem $u_k$.

Popravek za $u_k$ iščemo v podprostoru $u_k^\perp$. Korekcijsko enačbo

$$(I - u_k u_k^H)(A - \theta_k I)(I - u_k u_k^H)v = -r$$

rešimo le približno in z novim vektorjem razširimo bazo $\mathcal{V}_k$.

- Če za $v$ vzamemo $r$, dobimo Arnoldijevo metodo, saj je $\mathcal{L}(u_k, r) = \mathcal{L}(u_k, Au_k)$.
- Če $v$ aproksimiramo z $(D_A - \theta_k I)^{-1} r$, dobimo Davidsonovo metodo.
- Če točno rešimo sistem $(A - \theta_k I)v = -r$, dobimo $v = u_k$, kar ni v redu.

Pri J-D metodi iščemo približek $\widetilde{v}$ za rešitev sistema

$$(I - u_k u_k^H)(A - \theta_k I)(I - u_k u_k^H)v = -r, \quad v \perp u_k.$$

Če vzamemo vedno točno rešitev $v$, potem se izkaže, da imamo kvadratično konvergenco (pri simetričnih matrikah celo kubično), vendar moramo v tem primeru reševati sistem z matriko $A$.

Za približno reševanje sistema $(I - u_k u_k^H)(A - \theta_k I)(I - u_k u_k^H)v = -r$, $t \perp u_k$, naredimo nekaj korakov kakšne iterativne metode, npr. GMRES.

# Algoritem za Jacobi-Davidsonovo metodo

1. **Start:** Izberi začetni vektor $\|v_1\| = 1$.

   - Postavi $V_1 = [v_1]$, $u = v_1$, $\theta = u^H A u$ in izračunaj $r = Au - \theta u$.

2. **Zunanja zanka:** Za $k = 1, \ldots, m-1$:

   - Približno reši $(I - uu^H)(A - \theta I)(I - uu^H)v = -r, \quad v \perp u$.
   - Ortogonaliziraj $v$ glede na $V_k$ (RGS) in razširi $V_k$ v $V_{k+1}$.
   - Poišči dominantni lastni par $(\theta, s)$ matrike $V_{k+1}^H A V_{k+1}$, kjer je $\|s\| = 1$.
   - Izračunaj Ritzev vektor $u = V_{k+1}s$ in ostanek $r = Au - \theta u$.
   - Testiraj konvergenco in po potrebi končaj.

3. **Ponovni zagon:** Postavi $V_1 = [u]$ in ponovi točko 2.

# Obnašanje Jacobi-Davidsonove metode

V primeru simetrične matrike Ritzeve vrednosti monotono konvergirajo proti dominantni lastni vrednosti. V bližini rešitve imamo kubično konvergenco (če bi korekcijsko enačbo reševali točno). Ker namesto točnega reševanja uporabljamo GMRES (lahko pa tudi simetrično varianto MINRES), moramo poiskati kompromis med številom J-D iteracij in številom računanj produkta $Ax$.

V primeru nesimetrične matrike ali pri računanju notranjih lastnih vrednosti simetrične matrike imamo težave, ki pa se dajo odpraviti z uporabo harmoničnih Ritzevih vrednosti. Z njimi lahko poiščemo lastne vrednosti, ki so po absolutni vrednosti najmanjše in tako s premiki dobimo iskane lastne vrednosti.

# J-D kot nepopolna Newtonova metoda

Za skoraj vse vektorje $a, v$ je lastni vektor $x$ matrike $A$, skaliran tako, da je $a^H x = 1$, rešitev enačbe $F(x) = 0$, kjer je

$$F(x) = Ax - \theta x = 0, \quad \theta = \theta(x) = \frac{v^H A x}{v^H x}.$$

Če je $u_k$ približek za $x$, potem po Newtonovi metodi naslednji približek dobimo kot $u_{k+1} = u_k + t$, kjer je $t \perp a$,

$$JF(u_k)t = -r, \quad r = Au_k - \theta u_k$$

in $JF(x) = \left(I - \frac{xv^H}{v^H x}\right)(A - \theta I)$. Dobimo korekcijsko enačbo

$$\left(I - \frac{u_k v^H}{v^H u_k}\right)(A - \theta I)t = -r, \quad t \perp a.$$

# J-D kot nepopolna Newtonova metoda - 2

Za poljuben tak $y$, da je $a^H y \neq 0$, je sistem

$$\left(I - \frac{u_k v^H}{v^H u_k}\right)(A - \theta I)t = -r, \quad t \perp a,$$

ekivalenten sistemu

$$\left(I - \frac{u_k v^H}{v^H u_k}\right)(A - \theta I)\left(I - \frac{y a^H}{a^H y}\right)t = -r, \quad t \perp a.$$

Pri fiksnih $a, v, y$ imamo pri Newtonovi metodi kvadratično konvergenco, če pa $a, v, y$ konvergirajo k nekim vektorjem, imamo asimptotično kvadratično konvergenco. Če vzamemo $a = v = y = u_k$, dobimo korekcijsko enačbo J-D metode.

To pomeni, da lahko J-D obravnavamo kot nepopolno Newtonovo metodo. Če bi v vsakem koraku J-D metode točno rešili sistem, bi imeli asimptotično Newtonovo metodo, mi pa v vsakem koraku vzamemo le približek namesto točne rešitve.

# Predpogojevanje

Pri reševanje korekcijske enačbe je dobro uporabiti predpogojevanje.

$$(I - u_k u_k^H)(A - \theta_k I)(I - u_k u_k^H)v = -r$$

Denimo, da imamo na voljo matriko $K$, da je $K^{-1}(A - \theta_k I) \approx I$. Pri uporabi moramo tudi $K$ zožiti na isti podprostor, torej uporabimo

$$\widetilde{K} = (I - u_k u_k^H)K(I - u_k u_k^H).$$

Če za reševanje korekcijske enačbe uporabljamo metode podprostorov Krilova, moramo v vsakem koraku izračunati produkt

$$z = \widetilde{K}^{-1}\widetilde{A}w,$$

kjer je $\widetilde{A} = (I - u_k u_k^H)(A - \theta_k I)(I - u_k u_k^H)$.

1. $\widetilde{A}w = (I - u_k u_k^H)g$, kjer je $g = (A - \theta_k I)w$, saj je $u_k^H w = 0$.
2. reši $\widetilde{K}z = (1 - u_k u_k^H)g$, ker je $g \perp u_k$, je $Kz = g - \beta u_k$, $z = K^{-1}g - \beta K^{-1}u_k$, kjer je
$$\beta = \frac{u_k^H K^{-1} g}{u_k^H K^{-1} u_k}.$$

# Deflacija

Ko najdemo lastni par $(\lambda, x)$, nadaljujemo samo z podprostorom, ki ga razpenjajo preostali vektorji.

V primeru ortogonalne deflacije nadaljujemo z

$$(I - xx^H)A(I - xx^H).$$

V primeru več vektorjev nadaljujemo z

$$(I - ZZ^H)A(I - ZZ^H),$$

kjer je $AZ = ZS$ delni Schurov razcep. Stolpci matrike $Z$ so ortonormirani, matrika $S$ pa je zgornja trikotna z lastnimi vrednostmi na diagonali.

# Ponovni zagon (restart)

Ko je dimenzija iskalnega podprostora prevelika, izločimo manjši podprostor in nadaljujemo z njim.

V zadnjem koraku poiščemo Schurovo formo matrike $V_k^H A V_k$ in jo preuredimo tako, da so Ritzeve vrednosti, ki so blizu iskani tarči, v zgornjem delu. Potem nov začetni podprostor dobimo iz začetnih Schurovih vektorjev.

# Harmonične Ritzeve vrednosti

Če je $\mathcal{V}_k$ podprostor dimenzije $k$, potem je $\theta$ Ritzeva vrednost matrike $A$, če za nek $u \in \mathcal{V}_k$, $u \neq 0$, velja

$$Au - \theta u \perp \mathcal{V}_k.$$

$\theta$ je harmonična Ritzeva vrednost matrike $A$ glede na $\mathcal{W}_k$, če je $\theta^{-1}$ Ritzeva vrednost matrike $A^{-1}$ glede na $\mathcal{W}_k$.

Naj bo $\mathcal{V}_k$ podprostor dimenzije $k$. Potem je $\theta$ harmonična Ritzeva vrednost $A$ glede na $A\mathcal{V}_k$ natanko tedaj, ko za nek $u \in \mathcal{V}_k$, $u \neq 0$, velja

$$Au - \theta u \perp A\mathcal{V}_k.$$

Naj $V_k = [v_1 \ \cdots \ v_k]$ baza za $\mathcal{V}_k$ in $W_k = [w_1 \ \cdots \ w_k]$ baza za $\mathcal{W}_k := A\mathcal{V}_k$. Potem je $\theta$ harmonična Ritzeva vrednost $A$ glede na $A\mathcal{V}_k$ natanko tedaj, ko je

$$(W_k^H V_k)^{-1} W_k^H A V_k s = \theta s \quad \text{za nek } s \in \mathbb{C}^k, \ s \neq 0.$$

Če sta $V_k$ in $W_k$ taki ortogonalni bazi, da je $W_k = AV_k$, potem je $W_k^H A V_k = I$ in izračunati moramo lastne vrednosti matrike $(W_k^H V_k)^{-1}$. To lahko naredimo brez inverza, saj lastne vrednosti in vektorje dobimo iz lastnih parov $W_k^H V_k$.

## Korekcijska enačba za harmonične Ritzeve vrednosti

Z uporabo harmoničnih Ritzevih vrednosti lahko izračunamo lastne vrednosti z minimalno absolutno vrednostjo.

Naj bo $(\theta_k, u_k)$ harmonični Ritzev par v $k$-tem koraku J-D algoritma. Ostanek $r = Au_k - \theta u_k$ je ortogonalen na $z_k := Au_k$.

Pri korekcijski enačbi sedaj iščemo popravek, ki bo ortogonalen na $z_k$ (pri standardni metodi $u_k$). Namesto ortogonalne projekcije na $z_k^\perp$ uporabimo poševno projekcijo

$$I - \frac{u_k z_k^H}{z_k^H u_k},$$

saj tako $u_k$ pošljemo v $0$.

Nova korekcijska enačba je

$$\left(I - \frac{u_k z_k^H}{z_k^H u_k}\right)(A - \theta I)\left(I - \frac{u_k z_k^H}{z_k^H u_k}\right) t = -r.$$

# J-D algoritem s harmoničnimi Ritzevimi vrednostmi

**2. Zunanja zanka:** Za $k = 1, \ldots, m-1$:

- Približno reši

$$\left(I - \frac{uz^H}{z^H u}\right)(A - \theta I)\left(I - \frac{uz^H}{z^H u}\right) t = -r, \quad t \perp z.$$

- Ortogonaliziraj $t$ glede na $V_k$ (RGS) in razširi $V_k$ v $V_{k+1}$.
- Ortogonaliziraj $At$ glede na $W_k$ (RGS) in razširi $W_k$ v $W_{k+1}$.
- Poiščci minimalni lastni par $(\theta, s)$ matrike $(W_k^H V_k)^{-1} W_k^H A V_k$, kjer je $\|s\| = 1$.
- Izračunaj Ritzev vektor $u = V_{k+1}s$, ostanek $r = Au - \theta u$ in vektor $z = Au$.
- Testiraj konvergenco in po potrebi končaj.