

# O predmetu

Ukvarjali se bomo z

- iterativnim reševanjem linearnih sistemov, ki temelji na metodah podprostorov Krilova,
- metodami podprostorov za lastni problem lastnih vrednosti.

# Zgodovina

Z iterativnim reševanjem linearnih sistemov se je ukvarjal že Gauss. Leta 1823 je predlagal iterativno metodo za reševanje štirih enačb s štirimi neznankami, ki je izvirala iz kotnih meritev. Sistem je bil močno diagonalno dominanten.

Pri prvi metodi je zanemaril izvendiagonalne elemente in kot prvi približek za rešitev rešil le sistem z diagonalno matriko  $A$ , torej  $Dx_0 = b$ . To je potem iterativno popravljal v obliki  $x_{k+1} = x_k + y_k$ , kjer je  $Dy_k = b - Ax_k$ . Tu smo matriko  $A$  razdelili na spodnji trikotnik  $L$ , diagonalo  $D$  in zgornji trikotnik  $U$ . To ni nič drugega kot (Gauss-)Jacobijeva metoda, ki se imenuje še po nemškem matematiku in astronomu Jacobiju, ki je metodo uporabljal za računanje tirov planetov.

Jacobijeve rotacije, s katerimi lahko izračunamo lastne vrednosti simetrične matrike, je Jacobi uporabljal za to, da je matriko spravil v diagonalno dominantno obliko, potem pa je sistem lahko rešil z iterativnim postopkom.

Gauss je uporabljal tudi varianto, kjer je zanemaril zgornji del matrike in na začetku rešil sistem  $(L + D)x_0 = b$ , potem pa iterativno popravljal  $x_{k+1} = x_k + y_k$ , kjer je  $(L + D)y_k = b - Ax_k$ . To je Gauss-Seidelova metoda.

# Prednosti

Med prednostmi iterativnega reševanja linearnih sistemov je Gauss ugotovil, da za reševanje ne potrebuješ take koncentracije kot če sistem rešuješ preko eliminacij. V pismu prijatelju je bil navdušen, da te račune lahko opravljaš tudi, ko že napol spiš ali razmišljaš o drugih zadevah.

Za Gaussove majhne sisteme, ki so bili močno diagonalno dominantni, sta tako Jacobijeva kot Gauss-Seidelova metoda zelo hitro konvergirali. Za kakšne večje sisteme, ki se pojavljajo v današnjem času pa metodi nista več primerni in potrebujemo metode, ki hitreje konvergirajo.

Iterativne metode ponavadi uporabljamo na velikih in razpršenih sistemih. Pri takšnih matrikah lahko LU porabi preveč operacij in lahko tudi prostora. Dodatna prednost iterativnih metod je, da se lahko uporabnik sam odloči, s kakšno natančnostjo bo rešil sistem. Velikokrat ni potrebno rešiti sistem s polno natančnostjo in takrat so lahko iterativne metode konkurenčne LU razcepju tudi za polne matrike.

## 1.1 Osnovna ideja

Rešujemo linearni sistem  $Ax = b$ .

Namesto sistema z matriko  $A$ , rešimo enostavnejši sistem z matriko  $K$ :

$$Kx_0 = b$$

in vzamemo  $x_0$  kot približek za  $x$ .

To iterativno ponavljamo. Sedaj iščemo popravek  $z$ , da bo  $A(x_0 + z) = b$ , torej  $Az = b - Ax_0$ . Tako kot prej približek za popravek dobimo preko matrike  $K$ :

$$Kz_0 = b - Ax_0$$

in nadaljujemo z vektorjem  $x_1 = x_0 + z_0$ .

S ponavljanjem pridemo do iterativne metode.

Za matriko  $K$  ni nujno, da je ves čas fiksna.

## Osnovna iteracija

Za osnovno iteracijo tako dobimo

$$x_{k+1} = x_k + z_k = x_k + K^{-1}(b - Ax_k) = x_k + \tilde{b} - \tilde{A}x_k,$$

kjer smo označili  $\tilde{b} = K^{-1}b$  in  $\tilde{A} = K^{-1}A$ .

To si lahko predstavljamo kot osnovno iteracijo za predpogojeni linearni sistem  $\tilde{A}x = \tilde{b}$ , kjer je  $K = I$  aproksimacija za  $\tilde{A} = K^{-1}A$ .

Od zdaj naprej predpostavimo, da smo predpogojevanje že izvedli vnaprej in tako iteriramo  $Ax = b$  z aproksimacijo  $K = I$  za  $A$ .

Tako pridemo do *Richardsonove iteracije*

$$x_{k+1} = b + (I - A)x_k = x_k + r_k$$

za  $r_k = b - Ax_k$ .

## Richardsonova iteracija

Iz  $x_{k+1} = b + (I - A)x_k = x_k + r_k$ , kjer je  $r_k = b - Ax_k$ , dobimo

$$Ax_{k+1} = Ax_k + Ar_k$$

in

$$b - Ax_{k+1} = b - Ax_k - Ar_k,$$

torej

$$r_{k+1} = (I - A)r_k.$$

Od tod sledi

$$r_{k+1} = (I - A)^{k+1}r_0 = P_{k+1}(A)r_0,$$

kjer je  $P_{k+1}$  polinom stopnje  $k + 1$  z lastnostjo  $P_{k+1}(0) = 1$ .

To pomeni tudi

$$A(x - x_{k+1}) = P_{k+1}(A)A(x - x_0)$$

in v primeru nesingularne matrike  $A$  dobimo

$$x - x_{k+1} = P_{k+1}(A)(x - x_0).$$

# Konvergenca

Iz zveze  $r_{k+1} = (I - A)r_k$  sledi ocena  $\|r_{k+1}\| \leq \|I - A\| \cdot \|r_k\|$ .

Očitno imamo v primeru  $\|I - A\| < 1$  konvergenco za poljuben  $r_0$ .

Iz zveze  $r_{k+1} = P_{k+1}(A)r_0$  sledi, da lahko poljuben ostanek izrazimo kot polinom matrike  $A$  krat začetni ostanek. Denimo, da se da matrika  $A$  diagonalizirati, torej

$$Aw_j = \lambda_j w_j \quad \text{za } j = 1, \dots, n.$$

Sedaj lahko izrazimo  $r_0 = \sum_{j=1}^n \gamma_j w_j$  in dobimo

$$r_k = P_k(A)r_0 = \sum_{j=1}^n \gamma_j P_k(\lambda_j) w_j.$$

Hitrost konvergence je odvisna od tega, kako dobro polinom  $P_k$  zaduši lastne vrednosti matrike  $A$ .

## Prva posodobitev Richardsonove iteracije

Dodamo skalarje  $\alpha_k$  in iteriramo  $x_{k+1} = x_k + \alpha_k r_k$ . To pomeni

$$r_{k+1} = (I - \alpha_k A)r_k$$

in

$$P_k(A) = \prod_{j=1}^k (I - \alpha_j A).$$

Tako lahko pridemo do polinomov, ki bolje zadušijo lastne vrednosti. Za konverenco ni več nujno potrebno, da je  $\rho(I - A) < 1$ .

Npr., če so vse lastne vrednosti matrike  $A$  realne in na intervalu  $[a, b]$ , lahko  $\alpha_j$  izberemo kot ničle polinoma Čebiševa, preslikanega z  $[-1, 1]$  na  $[a, b]$ , ki ga skaliramo tako, da je  $P_k(0) = 1$ .

Tako pridemo do *iteracije Čebiševa*. Na prvi pogled je izbira  $\alpha_1, \dots, \alpha_k$  vezana na  $P_k$ , a preko tričlenske rekurzivne zveze pridemo do metode, ki deluje za poljuben  $k$ .



## 1.2 Druga posodobitev - podprostor Krilova

Drug pristop je, da shranimo vse približke  $x_0, x_1, \dots, x_k$  in iz linearnih kombinacij pridemo do še boljših približkov. Za  $x_{k+1}$  velja

$$x_{k+1} = r_0 + r_1 + \dots + r_k = \sum_{j=0}^k (I - A)^j r_0,$$

torej je  $x_{k+1} \in \text{Lin}(r_0, Ar_0, \dots, A^k r_0)$ .

**Definicija 1.** Za matriko  $A$  in vektor  $v$  definiramo podprostor Krilova  $\mathcal{K}_k(A, v)$  kot

$$\mathcal{K}_k(A, v) = \text{Lin}(v, Av, \dots, A^{k-1}v).$$

Zapišemo lahko tudi

$$\mathcal{K}_k(A, v) = \{p(A)v : p \in \mathcal{P}_{k-1}\},$$

kjer je  $\mathcal{P}_{k-1}$  prostor vseh polinomov stopnje manjše ali enake  $k$ .

$x_{k+1} \in \mathcal{K}_{k+1}(A, r_0)$ , torej  $x_{k+1} = Q_k(A)r_0$  za nek polinom  $Q_k$  stopnje  $k$ .

$r_{k+1} = b - Ax_{k+1} = (I - AQ_k(A))r_0 = \tilde{P}_{k+1}(A)r_0$ , kjer je  $\tilde{P}_{k+1}(0) = 1$ .

V podprostoru Krilova lahko pridemo do še boljših približkov kot z Richardsonovo iteracijo, kjer je  $r_{k+1} = (I - A)^{k+1}r_0$ .

## Kako dobimo približke iz podprostora Krilova

*Metode podprostorov Krilova* približke pridobivajo iz podprostora Krilova.

Rešujemo sistem  $Ax = b$ . Poznamo štiri pristope, kako metode podprostorov Krilova pridejo do približka  $x_k$  iz  $\mathcal{K}_k(A, r_0)$ .

1. *Ritz-Galerkinov pristop*: Izberemo  $x_k$ , pri katerem je ostanek  $b - Ax_k$  ortogonalen na  $\mathcal{K}_k(A, r_0)$ , torej  $b - Ax_k \perp \mathcal{K}_k(A, r_0)$ .  
Primer so CG (konjugirani gradienti), Lanczos, FOM, GENCG.
2. *Minimalni ostanek*: Izberemo  $x_k$ , pri katerem je norma ostanka  $\|b - Ax_k\|_2$  minimalna.  
Primer so GMRES, MINRES, ORTHODIR.
3. *Petrov-Galerkinov pristop*: Izberemo  $x_k$ , pri katerem je ostanek  $b - Ax_k$  ortogonalen na nek izbran  $k$ -razsežen podprostor.  
Primer so Bi-CG, QMR.
4. *Minimalna napaka*: Izberemo  $x_k$  iz podprostora  $A^T \mathcal{K}_k(A^T, r_0)$ , kjer je norma  $\|x - x_k\|_2$  minimalna.  
Primer so SYMMLQ, GMERR.

Poznamo tudi hibridne metode, kot so npr. CGS, Bi-CGSTAB in ostale.

Kaj izbrati?

## Baza za podprostor Krilova

Baza  $r_0, Ar_0, \dots, A^k r_0$  za  $\mathcal{K}_{k+1}(A, r_0)$  je zelo slabo pogojena in za praktično uporabo neprimerna. Potrebujemo ortonormirano bazo.

Naj bo  $u_j = A^{j-1} r_0$ . Potem so  $u_1, \dots, u_k$  bazni vektorji za  $\mathcal{K}_k(A, r_0)$ .

Iz vektorjev sestavimo matriko  $U_j = [u_1 \ u_2 \ \dots \ u_j]$ . Potem velja

$$AU_j = U_j B_j + u_{j+1} e_j^T,$$

kjer je  $B_j$  matrika velikosti  $j \times j$ , ki ima na glavni podiagonali enice, drugje pa povsod 0.

Iz  $Au_1 = u_2$ ,  $Au_2 = u_3$ ,  $Au_3 = u_4$  sledi

$$A [u_1 \ u_2 \ u_3] = [u_1 \ u_2 \ u_3] \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} + u_4 e_3^T,$$

oziroma

$$A [u_1 \ u_2 \ u_3] = [u_1 \ u_2 \ u_3] \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} + [0 \ 0 \ u_4].$$

## 1.3 Ortonormirana baza za podprostor Krilova

Imamo zvezo  $AU_k = U_k B_k + u_{k+1} e_k^T$ . Denimo, da za matriko  $U_k$  poznamo QR razcep

$$U_k = Q_k R_k,$$

kjer je  $Q_k^T Q_k = I$  in  $R_k$  zgornja trikotna. Sledi

$$A Q_k R_k = Q_k R_k B_k + u_{k+1} e_k^T,$$

od tod pa

$$\begin{aligned} A Q_k &= Q_k R_k B_k R_k^{-1} + u_{k+1} e_k^T R_k^{-1} \\ &= Q_k \tilde{H}_k + u_{k+1} e_k^T R_k^{-1} \\ &= Q_k \tilde{H}_k + \frac{1}{r_{kk}} u_{k+1} e_k^T, \end{aligned}$$

kjer smo definirali  $\tilde{H}_k = R_k B_k R_k^{-1}$ . Matrika  $\tilde{H}_k$  je zgornja Hessenbergova.

Podobno bi lahko zapisali razcep  $U_{k+1} = Q_{k+1} R_{k+1}$ , kjer je

$$R_{k+1} = \begin{bmatrix} R_k & \tilde{r} \\ 0 & r_{k+1,k+1} \end{bmatrix},$$

iz česar sledi  $u_{k+1} = Q_k \tilde{r} + r_{k+1,k+1} q_{k+1}$ .

## Ortonormirana baza za podprostor Krilova ...

Imamo zvezo  $AU_k = U_k B_k + u_{k+1} e_k^T$  in  $U_k = Q_k R_k$ . Dobimo

$$AQ_k = Q_k \tilde{H}_k + \frac{1}{r_{kk}} u_{k+1} e_k^T,$$

kjer je  $\tilde{H}_k = R_k B_k R_k^{-1}$  zgornja Hessenbergova matrika. Iz  $U_{k+1} = Q_{k+1} R_{k+1}$ , kjer je

$$R_{k+1} = \begin{bmatrix} R_k & \tilde{r} \\ 0 & r_{k+1,k+1} \end{bmatrix},$$

sledi  $u_{k+1} = Q_k \tilde{r} + r_{k+1,k+1} q_{k+1}$ .

To združimo v

$$\begin{aligned} AQ_k &= Q_k \left( \tilde{H}_k + \frac{1}{r_{kk}} \tilde{r} e_k^T \right) + \frac{r_{k+1,k+1}}{r_{kk}} q_{k+1} e_k^T \\ &= Q_k H_k + \alpha q_{k+1} e_k^T. \end{aligned}$$

Od tod vidimo, da je  $Q_k^T A Q_k = H_k$ , kjer je  $H_k$  zgornja Hessenbergova in

$$q_{k+1}^T A q_k = \alpha,$$

od koder iz  $Q_{k+1}^T A Q_{k+1} = H_{k+1}$  sledi  $\alpha = h_{k+1,k}$ .