0

Metoda konjugovaných gradientů

Luděk Kučera
MFF UK

11. ledna 2017

V tomto textu je popsáno, jak metodou konjugovaných gradientů řešit soustavu lineárních rovnic $\mathbf{Ax} = \mathbf{b}$, kde $\mathbf{b}$ je daný vektor a $\mathbf{A}$ je symetrická a pozitivně definitní reálná matice. Jedná se o metodu iterativní: vycházejíce z libovolného bodu $\mathbf{x}_0$ v příslušném Euklidovském prostoru, budeme postupně procházet (opakováním stále stejného výpočtu) body $\mathbf{x}_1, \mathbf{x}_2, \ldots$, až se nakonec dostaneme do bodu, který je řešením soustavy, nebo alespoň tak blízko k němu, že nám daná hodnota bude vyhovovat jako přibližné řešení soustavy.

1 Řešení soustavy lineárních rovnic minimalizací kvadratického funkcionálu

Definujme funkcionál f takto:

$$f(x) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{x}^T\mathbf{b} \qquad \text{pro každé } \mathbf{x} \in E_n.$$

Lemma 1 *Řešení $\mathbf{x}_{sol}$ soustavy $\mathbf{Ax} = \mathbf{b}$ je jediné minimum funkcionálu f.*

Důkaz: Uvažujme vektor $\mathbf{x} \in E_n$ a řešení $\mathbf{x}_{sol}$ soustavy $\mathbf{Ax} = \mathbf{b}$ (*sol* jako 'solution', anglicky 'řešení'). Položme $\mathbf{e} = \mathbf{x} - \mathbf{x}_{sol}$. Pak

$$f(\mathbf{x}) = f(\mathbf{x}_{sol} + \mathbf{e}) = \frac{1}{2}(\mathbf{x}_{sol} + \mathbf{e})^T\mathbf{A}(\mathbf{x}_{sol} + \mathbf{e}) - (\mathbf{x}_{sol} + \mathbf{e})^T\mathbf{b} =$$

$$= \frac{1}{2}(\mathbf{x}_{sol}^T\mathbf{A}\mathbf{x}_{sol} + \mathbf{x}_{sol}^T\mathbf{A}\mathbf{e} + \mathbf{e}^T\mathbf{A}\mathbf{x}_{sol} + \mathbf{e}^T\mathbf{A}\mathbf{e}) - \mathbf{x}_{sol}^T\mathbf{b} - \mathbf{e}^T\mathbf{b} =$$

$$= (\frac{1}{2}\mathbf{x}_{sol}^T\mathbf{A}\mathbf{x}_{sol} - \mathbf{x}_{sol}^T\mathbf{b}) + \mathbf{e}^T\mathbf{A}\mathbf{e} + \mathbf{e}^T(\mathbf{A}\mathbf{x}_{sol} - \mathbf{b}) = f(\mathbf{x}_{sol}) + \mathbf{e}^T\mathbf{A}\mathbf{e},$$

protože $\mathbf{x}_{sol}^T\mathbf{A}\mathbf{e} = \mathbf{e}^T\mathbf{A}\mathbf{x}_{sol}$ a proto $\frac{1}{2}(\mathbf{x}_{sol}^T\mathbf{A}\mathbf{e} + \mathbf{e}^T\mathbf{A}\mathbf{x}_{sol}) = \mathbf{e}^T\mathbf{A}\mathbf{x}_{sol}$.

Jelikož je $\mathbf{A}$ pozitivně definitní, je hodnota výrazu $\mathbf{e}^T\mathbf{A}\mathbf{e}$ vždy nezáporná a je rovna nule (své nejmenší možné hodnotě) právě když $\mathbf{e} = \mathbf{0}$, neboli když $\mathbf{x} = \mathbf{x}_{sol}$. ♣

2 Minimalizace funkcionálu v daném směru

Předpokládejme nyní, že se nacházíme v bodě $\mathbf{x}$ a vydáme se z něj ve směru vektoru $\mathbf{p}$ (vpřed nebo i vzad). Navštívíme tedy body $\mathbf{x} + \alpha\mathbf{p}$, kde α je reálné číslo. Naším cílem je určit, kde na této dráze má funkcionál f nejmenší hodnotu. Hodnota funkcionálu f na této dráze je

$$f(\mathbf{x} + \alpha\mathbf{p}) = \frac{1}{2}(\mathbf{x} + \alpha\mathbf{p})^T\mathbf{A}(\mathbf{x} + \alpha\mathbf{p}) - (\mathbf{x} + \alpha\mathbf{p})^T\mathbf{b} =$$

$$= \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} + \alpha\mathbf{p}^T\mathbf{A}\mathbf{x} + \frac{1}{2}\alpha^2\mathbf{p}^T\mathbf{A}\mathbf{p} - \mathbf{x}^T\mathbf{b} - \alpha\mathbf{p}^T b = U + \alpha V + \alpha^2 W,$$

kde $U = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{x}^T\mathbf{b}$, $V = \mathbf{p}^T\mathbf{A}\mathbf{x} - \mathbf{p}^T\mathbf{b}$, $W = \frac{1}{2}\mathbf{p}^T\mathbf{A}\mathbf{p}$.

Pro pevné x a p jsou U, V a W konstanty a výraz pro $f(\mathbf{x} + \alpha\mathbf{p})$ je kvadratickou funkcí α. Je známo, že tato funkce má minimum, když se její derivace podle α rovná 0. Ta derivace je obecně rovna $2W\alpha + V$ a tedy minimum nabývá pro $\alpha = -\frac{V}{2W}$, neboli pro

$$\alpha = \frac{\mathbf{p}^T(\mathbf{b} - \mathbf{A}\mathbf{x})}{\mathbf{p}^T\mathbf{A}\mathbf{p}}.$$

Mimimum na dráze tvořené body $\mathbf{x} + \alpha\mathbf{p}$ je v bodě

$$\mathbf{x} + \frac{\mathbf{p}^T(\mathbf{b} - \mathbf{A}\mathbf{x})}{\mathbf{p}^T\mathbf{A}\mathbf{p}}\mathbf{p} \tag{1}$$

3 Směr největšího spádu

Dále si představme, že jsme v bodě $\mathbf{x}$ a chceme se pustit směrem nejprudšího klesání funkcionálu f. Který to je směr?

Vezměme vektor $\mathbf{e}$ délky 1 a vydejme se v jeho směru z bodu $\mathbf{x}$. Položme $\varphi(\alpha) = f(\mathbf{x} + \alpha\mathbf{e})$. Pro pevné $\mathbf{e}$ je rychlost změny funkcionálu f v bodě $\mathbf{x}$ při pohybu ve směru $\mathbf{e}$ rovna hodnotě derivace funkce φ podle α v bodě $\alpha = 0$. Propočítejme to:

$$\varphi(\alpha) = f(\mathbf{x} + \alpha\mathbf{e}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} + \alpha\mathbf{e}^T\mathbf{A}\mathbf{x} + \alpha^2\mathbf{e}^T\mathbf{A}\mathbf{e} - \mathbf{x}^T\mathbf{b} - \alpha\mathbf{e}^T\mathbf{b}$$

$$\frac{\partial\varphi(\alpha)}{\partial\alpha} = \mathbf{e}^T\mathbf{A}\mathbf{x} + 2\alpha\mathbf{e}^T\mathbf{A}\mathbf{e} - \mathbf{e}^T\mathbf{b}$$

a tedy hodnota této derivace pro $\alpha = 0$ je

$$\left.\frac{\partial\varphi(\alpha)}{\partial\alpha}\right|_{\alpha=0} = \mathbf{e}^T\mathbf{A}\mathbf{x} - \mathbf{e}^T\mathbf{b} = -\mathbf{e}^T(\mathbf{b} - \mathbf{A}\mathbf{x}).$$

Nyní se zamysleme nad tím, v jakém směru je ta derivace nejmenší, tedy skalární součin $\mathbf{e}^T(\mathbf{b} - \mathbf{A}\mathbf{x})$ největší. Jasně, je to když $\mathbf{e}$ jde ve směru $\mathbf{b} - \mathbf{A}\mathbf{x}$.

4 Metoda největšího spádu

Máme-li tedy nyní řešit rovnici $\mathbf{Ax} = \mathbf{b}$, stačí nalézt minimum funkcionálu $f(x) = \frac{1}{2}\mathbf{x}^T\mathbf{Ax} - \mathbf{x}^T\mathbf{b}$, které můžeme hledat iterativně takto: postavme se do libovolného bodu, který označíme $\mathbf{x}_0$. Často se volí $\mathbf{x}_0 = \mathbf{0}$, nemáme-li nějaký dobrý důvod zvolit jiný bod. Vydejme se z tohoto bodu směrem největšího klesání funkcionálu f, tedy ve směru daném vektorem $\mathbf{r}_0 = \mathbf{b} - \mathbf{Ax}_0$. V tomto směru se posuneme tak hluboko (vzhledem k f), jak to jen jde. O kolikanásobek vektoru r se posuneme, je popsáno výše.

Postupem z předchozího odstavce se dostaneme do bodu, který označíme $\mathbf{x}_1$ a postup stále opakujeme. Jsme-li tedy obecně v bodě $\mathbf{x}_i$, položíme $\mathbf{r}_i = \mathbf{b} - \mathbf{Ax}_i$ (povšimněte si, že $\mathbf{r}_i$ nám také udává, jak jsme daleko s výpočtem, jak je $\mathbf{Ax}_i$ daleko od vektoru $\mathbf{b}$, kterému by se nakonec mělo rovnat), vydáme se ve směru r_i, dojdeme v tom směru do místa, ve kterém je funkcionál f nejnižší a to označíme jako $\mathbf{x}_{i+1}$, atd. Pokračujeme tak dlouho, dokud není $\mathbf{r}_i$ (kterému se říkává i-té *residuum*, odtud označení $\mathbf{r}_i$) rovno nule (nalezeno řešení) nebo alespoň dostatečně malé (nalezeno dostatečně kvalitní přibližné řešení).

Této metodě říkáme *metoda nejprudšího spádu* (anglicky "steepest descent").

5 Metoda konjugovaných gradientů

V roce 1952 si Magnus Hestenes a Eduard Stiefel povšimli, že v elipsoidové geometrii pozitivně definitních matic je být výhodnější postupovat trochu jinak a lépe, než metodou nejprudšího spádu: způsobem, kterému se nyní říká metoda *konjugovaných gradientů* (anglicky "conjugate gradients", odtud i zkratka CG).

Je dobře známo, že je-li $\mathbf{A}$ pozitivně definitní a symetrická reálná matice, pak vzorec $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{Ay}$ má vlastnosti skalárního součinu, tedy $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ s rovností pouze pro $\mathbf{x} = \mathbf{0}$, operace je symetrická a $\langle \alpha\mathbf{x}, \mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle$. Řekneme, že dva vektory $\mathbf{x}$ a $\mathbf{y}$ jsou *konjugované*, pokud platí $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ (tedy jsou na sebe "kolmé" vzhledem ke skalárnímu součinu, určenému maticí $\mathbf{A}$).

Metoda konjugovaných gradientů postupuje velmi podobně jako metoda nejprudšího spádu, ale budeme se snažit, aby směry, ve kterých se v jednotlivých iteracích pohybujeme (gradienty), byly navzájem konjugované.

Postup začíná v libovolně zvoleném vektoru $\mathbf{x}_0$ a vytváří posloupnost $\mathbf{x}_0, \mathbf{r}_0, \mathbf{p}_k, \mathbf{x}_1, \mathbf{r}_1, \mathbf{p}_1, \ldots$ následujícím způsobem:

Metoda konjugovaných gradientů (neupravená)
Vstupní data:
celé číslo n (dimenze)
symetrická pozitivně definitní matice $\mathbf{A}$ velikosti $n \times n$,
vektor $\mathbf{b}$ dimenze n (pravá strana soustavy),
počáteční vektor $\mathbf{x}_0$.

pro $k = 0, 1, 2, \ldots$ {

$$\mathbf{r}_k = \mathbf{b} - \mathbf{A}\mathbf{x}_k \ ; \quad \text{pokud } \mathbf{r}_k = \mathbf{0}\text{, ukonči výpočet;} \tag{2}$$

$$\mathbf{p}_k = \mathbf{r}_k - \sum_{j=0}^{k-1} \frac{\mathbf{r}_k^T \mathbf{A}\mathbf{p}_j}{\mathbf{p}_j^T \mathbf{A}\mathbf{p}_j} \mathbf{p}_j \ ; \tag{3}$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{\mathbf{p}_k^T \mathbf{r}_k}{\mathbf{p}_k^T \mathbf{A}\mathbf{p}_k} \mathbf{p}_k \ ; \tag{4}$$

}

První krok se udělá stejně jako v metodě nejprudšího spádu; vyrazíme směrem $\mathbf{p}_0$, který bude roven směru $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$. Když se ale dostaneme po k iteracích do bodu $\mathbf{x}_k$, nebudeme dál postupovat nejprudším spádem přesně ve směru residua $\mathbf{r}_k = \mathbf{b} - \mathbf{A}\mathbf{x}_k$, ale ve směru $\mathbf{p}_k$, který vznikne tak, že si směr $\mathbf{r}_k$ upravíme tak, aby byl *konjugovaný* ke všem směrům $\mathbf{p}_0, \ldots, \mathbf{p}_{k-1}$, ve kterých jsme se pohybovali dříve.

Tento pozměněný směr $\mathbf{p}_k$ ($\mathbf{p}$ od anglického 'progress') dostaneme z $\mathbf{r}_k$ podobně jako při Gramm-Schmidtově ortogonalizaci tak, že od residua $\mathbf{r}_k$ odečteme jeho složky v předchozích směrech $\mathbf{p}_0 = \mathbf{r}_0, \mathbf{p}_1, \ldots, \mathbf{p}_{k-1}$, ale místo obvyklé kolmosti používáme "kolmost" ve smyslu skalárního součinu definovaného maticí $\mathbf{A}$, tedy konjugovanost.

Pro každé $j < k$ tedy chceme od vektoru $\mathbf{r}_k$ odečíst takový násobek α_j vektoru $\mathbf{p}_j$, aby výsledek odečítání byl konjugovaný s $\mathbf{p}_j$. Pro hodnotu násobku α_j tedy máme následující rovnici: $(\mathbf{r}_k - \alpha_j \mathbf{p}_j)^T \mathbf{A}\mathbf{p}_j = \mathbf{0}$, neboli $\mathbf{r}_k^T \mathbf{A}\mathbf{p}_j = \alpha_j \mathbf{p}_j^T \mathbf{A}\mathbf{p}_j$, z čehož dostáváme

$$\alpha_j = \frac{\mathbf{r}_k^T \mathbf{A}\mathbf{p}_j}{\mathbf{p}_j^T \mathbf{A}\mathbf{p}_j}.$$

To nám tedy dává vzorec (3) pro $\mathbf{p}_k$, uvedený výše.

Jestliže je takto vzniklé $\mathbf{p}_k$ nenulový vektor, přejdeme z $\mathbf{x}_k$ ve směru daném vektorem $\mathbf{p}_k$ do bodu $\mathbf{x}_{k+1}$, který minimalizuje funkcionál f na této přímce. Na základě (1) tedy se dostaneme ke vzorci (4) pro x_{k+1}, uvedenému výše.

Pokud by nastal případ, že $\mathbf{r}_k$ je nulový vektor, výpočet se přeruší, protože v takovém případě $\mathbf{x}_k$ je řešení soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$ a není nutno dále pokračovat.

Potíž při výpočtech by mohla nastat, pokud by se stalo, že $\mathbf{p}_k = 0$, protože hned v následujícím výpočtu $\mathbf{x}_{k+1}$ ve vzorci (4) by došlo k dělení nulou. Za chvíli ale bude dokázáno, že k tomu nikdy nemůže dojít.

Nyní dokážeme tři jednoduchá lemmata, která budou mít řadu velmi důležitých důsledků.

Lemma 2 *Pokud je $\mathbf{p}_k$ definováno, pak je lineární kombinací vektorů $\mathbf{r}_0, \ldots, \mathbf{r}_k$.*

Důkaz: Lemma dokážeme indukcí podle k. Pro $k = 0$ tvrzení platí, protože $\mathbf{p}_0 = \mathbf{r}_0$. A jestliže platí pro $k - 1$, pak z (3) okamžitě plyne, že platí také pro k. ♣

Lemma 3 *Pokud je $\mathbf{r}_{k+1}$ definováno, pak*

$$\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{A}\mathbf{p}_k, \quad \textit{kde} \quad \alpha_k = \frac{\mathbf{p}_k^T \mathbf{r}_k}{\mathbf{p}_k^T \mathbf{A}\mathbf{p}_k}.$$

Důkaz: Na základě (4) je

$$\mathbf{x}_{k+1} - \mathbf{x}_k = \frac{\mathbf{p}_k^T \mathbf{r}_k}{\mathbf{p}_k^T \mathbf{A}\mathbf{p}_k} \mathbf{p}_k,$$

takže

$$\mathbf{r}_{k+1} - \mathbf{r}_k = (\mathbf{b} - \mathbf{A}\mathbf{x}_{k+1}) - (\mathbf{b} - \mathbf{A}\mathbf{x}_k) = -(\mathbf{A}\mathbf{x}_{k+1} - \mathbf{A}\mathbf{x}_k) = -\frac{\mathbf{p}_k^T \mathbf{r}_k}{\mathbf{p}_k^T \mathbf{A}\mathbf{p}_k} \mathbf{A}\mathbf{p}_k.$$

♣

Lemma 4 (Rezidua jsou na sebe kolmá)
Nechť pro nějaké celé nezáporné číslo k platí, že $\mathbf{x}_j$ a $\mathbf{r}_j$ jsou definována a $\mathbf{r}_j \neq \mathbf{0}$ pro všechna $j = 0, 1, \ldots, k$ a navíc $\mathbf{r}_i^T \mathbf{r}_j = 0$ pro všechna celá i a j taková, že $0 \leq i < j \leq k$.
Potom $\mathbf{p}_k \neq \mathbf{0}$, také $\mathbf{x}_{k+1}$ a $\mathbf{r}_{k+1}$ jsou řádně definovány a $\mathbf{r}_{k+1}^T \mathbf{r}_j = 0$ pro $j = 0, 1, \ldots, k$.

Důkaz: Jelikož jsou rezidua $\mathbf{r}_0, \ldots, \mathbf{r}_k$ nenulová a navzájem kolmá, jsou také lineárně nezávislá. Suma v rovnici (3) je lineární kombinací vektorů $\mathbf{p}_0, \ldots, \mathbf{p}_{k-1}$ a tedy na základě předchozího lemmatu také lineární kombinací vektorů $\mathbf{r}_0, \ldots, \mathbf{r}_{k-1}$. Podle (3) je tedy p_k *netriviální* lineární kombinací vektorů $\mathbf{r}_0, \ldots, \mathbf{r}_k$. Proto $\mathbf{p}_k$ nemůže být nulový vektor. A je-li $\mathbf{p}_k \neq \mathbf{0}$, pak jsou také definovány vektory $\mathbf{x}_{k+1}$ a $\mathbf{r}_{k+1}$.

Nakonec spočítejme hodnotu $\mathbf{r}_{k+1}^T \mathbf{r}_j$ pro zvolené celé j v rozmezí $0 \leq j \leq k$:

$$\mathbf{r}_{k+1}^T \mathbf{r}_j = (\mathbf{r}_k - \alpha_k \mathbf{A}\mathbf{p}_k)^T \mathbf{r}_j = \mathbf{r}_k^T \mathbf{r}_j - \alpha_k \mathbf{p}_k^T \mathbf{A} \left(\mathbf{p}_j + \sum_{i=0}^{j-1} \frac{\mathbf{r}_j^T \mathbf{A}\mathbf{p}_i}{\mathbf{p}_i^T \mathbf{A}\mathbf{p}_i} \mathbf{p}_i \right) = \mathbf{r}_k^T \mathbf{r}_j - \alpha_k \mathbf{p}_k^T \mathbf{A}\mathbf{p}_j,$$

kde

$$\alpha_k = \frac{\mathbf{p}_k^T \mathbf{r}_k}{\mathbf{p}_k^T \mathbf{A} \mathbf{p}_k},$$

protože $\mathbf{p}_k$ bylo zvoleno tak, aby platilo $\mathbf{p}_k^T \mathbf{A} \mathbf{p}_i = 0$ pro $i < k$.

Víme nyní, že pro $j < k$

$$\mathbf{p}_j^T \mathbf{r}_k = r_j - \sum_{i=0}^{j-1} \frac{\mathbf{r}_j^T \mathbf{A} \mathbf{p}_i}{\mathbf{p}_i^T \mathbf{A} \mathbf{p}_i} \mathbf{p}_i = \mathbf{r}_j \mathbf{r}_k,$$

a

$$\mathbf{p}_k^T \mathbf{r}_k = (\mathbf{r}_k - \sum_{j=0}^{k-1} \frac{\mathbf{r}_k^T \mathbf{A} \mathbf{p}_j}{\mathbf{p}_j^T \mathbf{A} \mathbf{p}_j} \mathbf{p}_j)^T \mathbf{r}_k = \mathbf{r}_k^T \mathbf{r}_k.$$

Jestliže nyní $j < k$, pak z předpokladů věty plyne, že $\mathbf{r}_k^T \mathbf{r}_j = 0$ a $\mathbf{p}_k^T \mathbf{A} \mathbf{p}_j = 0$, takže i $\mathbf{r}_{k+1}^T \mathbf{r}_j = 0$. Pokud naopak $j = k$, pak

$$\mathbf{r}_k^T \mathbf{r}_j - \alpha_k \mathbf{p}_k^T \mathbf{A} \mathbf{p}_j = \mathbf{r}_k^T \mathbf{r}_k - \mathbf{p}_k^T \mathbf{r}^k = \mathbf{r}_k^T \mathbf{r}_k - \mathbf{r}_k^T \mathbf{r}^k = 0.$$

♣

Lemma říká, že residua jsou na sebe kolmá v klasickém smyslu (obvyklý skalární součin), nikoli konjugovaná (skalární součin indukovaný maticí $\mathbf{A}$).

Jední z velmi významných důsledků lemmatu je následující věta (je třeba poznamenat, že platí za předpokladu, že všechny výpočty se provádějí absolutně přesně, bez zaokrouhlovacích chyb):

Věta 1 *Metoda konjugovaných gradientů nalezne přesné řešení soustavy rovnic* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *po nejvýše* n *krocích.*

Důkaz: Kdyby se metoda konjungovaných gradientů nezastavila po n krocích, vytvořila by nenulové vektory $\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_n$, které by ale podle předchozího lemmatu byly na sebe kolmé, tedy lineárně nezávislé, což však v prostoru dimenze n není možné. ♣

Další velmi užitečné lemma říká, že ve vzorci (3) pro výpočet $\mathbf{p}_k$ jsou všechny sčítance v sumě s výjimkou sčítance pro $j = k - 1$ rovny nule, takže je tento vzorec možno velice zjednodušit:

Lemma 5

$$\mathbf{r}_k^T \mathbf{A} \mathbf{p}_j = 0 \qquad \textit{pro} \quad j = 0, \ldots, k-2$$

$$\frac{\mathbf{r}_{k+1}^T \mathbf{A} \mathbf{p}_k}{\mathbf{p}_k^T \mathbf{A} \mathbf{p}_k} = \frac{\mathbf{r}_{k+1}^T \mathbf{r}_{k+1}}{\mathbf{r}_k^T \mathbf{r}_k}$$

Důkaz: Na základě lemmatu 3**0** platí, že

$$\mathbf{r}_k^T \mathbf{A} \mathbf{p}_j = \mathbf{r}_k^T (-\frac{1}{\alpha_{j-1}} (\mathbf{r}_j - \mathbf{r}_{j-1}) = 0 \qquad \text{pro } j = 0, 1, \ldots, k-2$$

$$\frac{\mathbf{r}_{k+1}^T \mathbf{A} \mathbf{p}_k}{\mathbf{p}_k^T \mathbf{A} \mathbf{p}_k} = \frac{\mathbf{r}_{k+1}^T \mathbf{A} \mathbf{p}_{k-1}}{\alpha_{k-1} \mathbf{p}_{k-1}^T \mathbf{A} \mathbf{p}_{k-1}} = \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{r}_{k-1}^T \mathbf{r}_{k-1}}$$

♣

Důsledek 1

$$\mathbf{p}_{k+1} = \mathbf{r}_{k+1} + \beta_k \mathbf{p}_k, \qquad kde\ \beta_k = \frac{\mathbf{r}_{k+1}^T \mathbf{r}_{k+1}}{\mathbf{r}_k^T \mathbf{r}_k}$$

Důkaz: ♣

Metoda konjugovaných gradientů (s indexy)
Vstupní data:
celé číslo n (dimenze)
symetrická pozitivně definitní matice $\mathbf{A}$ velikosti $n \times n$,
vektor $\mathbf{b}$ dimenze n (pravá strana soustavy),
počáteční vektor $\mathbf{x}_0$.

$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$;
$\mathbf{p}_0 = \mathbf{r}_0$;
pro $k = 0, 1, 2, \ldots$ {

$$\begin{aligned}
\mathbf{z}_k &= \mathbf{A}\mathbf{p}_k \ ; \\
\alpha_k &= \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{p}_k^T \mathbf{z}_k} \ ; \\
\mathbf{x}_{k+1} &= \mathbf{x}_k + \alpha_k \mathbf{p}_k \ ; \\
\mathbf{r}_{k+1} &= \mathbf{r}_k - \alpha_k \mathbf{z}_k; \quad \text{pokud } \mathbf{r}_k = \mathbf{0}\text{, ukonči výpočet;} \\
\beta_k &= \frac{\mathbf{r}_{k+1}^T \mathbf{r}_{k+1}}{\mathbf{r}_k^T \mathbf{r}_k} \ ; \\
\mathbf{p}_{k+1} &= \mathbf{r}_{k+1} + \beta_k \mathbf{p}_k \ ;
\end{aligned}$$

}

Metoda konjugovaných gradientů (konečný tvar)
Vstupní data:
celé číslo n (dimenze)
symetrická pozitivně definitní matice $\mathbf{A}$ velikosti $n \times n$,
vektor $\mathbf{b}$ dimenze n (pravá strana soustavy),
počáteční vektor $\mathbf{x}_0$.

$\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x}_0$;
$lrr = \mathbf{r}^T\mathbf{r}$;
$\mathbf{p} = \mathbf{r}$;
pro $k = 0, 1, 2, \ldots$ {

$$\begin{aligned}
\mathbf{z} &= \mathbf{A}\mathbf{p} \ ; \\
\alpha &= \frac{lrr}{\mathbf{p}^T\mathbf{z}} \ ; \\
\mathbf{x} &= \mathbf{x} + \alpha\mathbf{p} \ ; \\
\mathbf{r} &= \mathbf{r} - \alpha\mathbf{z}; \quad \text{pokud } \mathbf{r} = \mathbf{0}\text{, ukonči výpočet;} \\
rr &= \mathbf{r}^T\mathbf{r} \ ; \\
\beta &= \frac{rr}{lrr} \ ; \\
lrr &= rr \\
\mathbf{p} &= \mathbf{r} + \beta\mathbf{p} \ ;
\end{aligned}$$

}