Metoda konjugiranih gradijenata

1 Uvod u problem

Zadani su simetrična i pozitivno definitna kvadratna matrica $\mathbf{A} \in \mathbb{R}^{n \times n}$, realni vektor $\mathbf{b} \in \mathbb{R}^n$ te $c \in \mathbb{R}$. Definirajmo funkciju $f : \mathbb{R}^n \to \mathbb{R}$ formulom $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{b}^T\mathbf{x} + c$ te promatrajmo sljedeći problem:

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{b}^T\mathbf{x} + c \to \min_{\mathbf{x}\in\mathbb{R}^n}. \tag{1}$$

Neka je $\mathbf{x}^* = \operatorname{argmin}_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x})$ rješenje problema (1). Uočimo da je

$$\nabla f(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b} \text{ te } \nabla^2 f(\mathbf{x}) = \mathbf{A}.$$

Kako je prema pretpostavci matrica $\mathbf{A}$ pozitivno definitna, te stoga i regularna (vidi Domaću zadaću), vektor $\mathbf{x}^*$ jedinstveno rješenje sustava linearnih jednadžbi

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \tag{2}$$

To znači da je problem minimizacije (1) ekvivalentan problemu rješavanju sustava linearnih jednadžbi (2).

Neka su $\alpha \in \mathbb{R}$ i $\mathbf{p} \in \mathbb{R}^n$. Razvojem izraza $f(\mathbf{x} + \alpha\mathbf{p})$ u Taylorov red u okolini točke $\mathbf{x}$ dobivamo

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

Minimizacijski problem (1) rješavat ćemo iterativnom metodom. Pretpostavimo u tu svrhu da je $\mathbf{x}_j$ neka aproskimacija rješenja $\mathbf{x}^*$. Novu aproksimaciju $\mathbf{x}_{j+1}$ tražit ćemo u obliku $\mathbf{x}_{j+1} = \mathbf{x_j} + \alpha_j\mathbf{p}_j$, pri čemu je $\mathbf{p}_j$ vektor smjera, a α_j odgovarajuća duljina koraka. Ako nam je poznat vektor smjera $\mathbf{p}_j$ onda duljinu optimalnu duljinu koraka α_j možemo dobiti rješavanjem problema jednodimenzionalne minimizacije

$$f(\mathbf{x}_j + \alpha\mathbf{p}_j) \to \min_{\alpha\in\mathbb{R}}.$$

Imamo

$$\begin{aligned} \alpha_j &= \operatorname{argmin}_{\alpha\in\mathbb{R}} f(\mathbf{x}_j + \alpha\mathbf{p}_j) \\ &= \operatorname{argmin}_{\alpha\in\mathbb{R}} \left(f(\mathbf{x}_j) + \alpha\mathbf{p}_j^T(\mathbf{A}\mathbf{x} - \mathbf{b}) + \frac{1}{2}\alpha^2\mathbf{p}^T\mathbf{A}\mathbf{p} \right). \end{aligned}$$

Funkcija $\alpha \mapsto f(\mathbf{x}_j) + \alpha\mathbf{p}_j^T(\mathbf{A}\mathbf{x}_j - \mathbf{b}) + \frac{1}{2}\alpha^2\mathbf{p}_j^T\mathbf{A}\mathbf{p}_j$ je realna kvadratna funkcija, koja postiže jedinstveni globalni minimum u točki

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

1.0.1 Gradijentna metoda

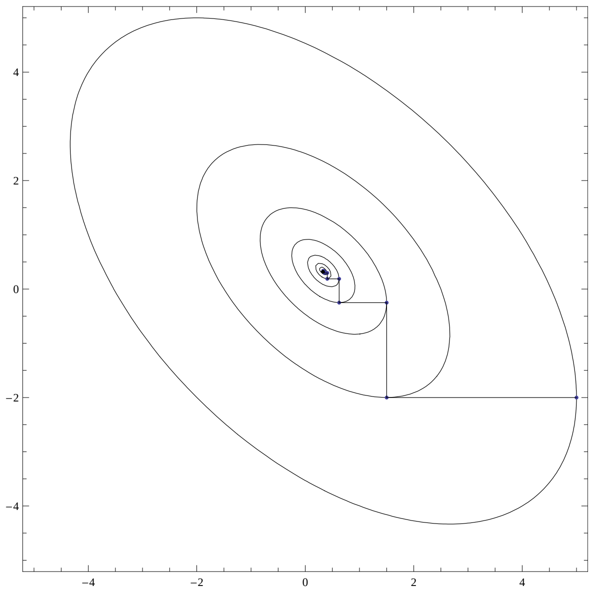
Specijalno u gradijentnoj metodi za vektor smjera $\mathbf{p}_j$ uzimamo $\mathbf{p}_j = -\nabla f(\mathbf{x}_j) = -\mathbf{A}\mathbf{x}_j + \mathbf{b}$, odakle je

$$\alpha_j^* = \frac{\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2}{(\mathbf{b} - \mathbf{A}\mathbf{x}_j)^T \mathbf{A}(\mathbf{b} - \mathbf{A}\mathbf{x}_j)} = \frac{\|\mathbf{p}_j\|_2^2}{\mathbf{p}_j^T \mathbf{A}\mathbf{p}_j}.$$

Za vektor smjera $\mathbf{p}_{j+1}$ vrijedi

$$\begin{aligned} \mathbf{p}_{j+1} &= -\nabla f(\mathbf{x}_{j+1}) = \mathbf{b} - \mathbf{A}\mathbf{x}_{j+1} \\ &= \mathbf{b} - \mathbf{A}(x_j + \alpha_j \mathbf{p_j}) \\ &= \mathbf{p}_j - \alpha_j \mathbf{A}\mathbf{p}_j, \end{aligned}$$

odakle je



Slika 1: Vektori smjerova gradijentne metode kroz 17 iteracija

$$\mathbf{p}_{j+1}^T \mathbf{p}_j = \|\mathbf{p}_j\|_2^2 - \alpha_j \mathbf{p_j}^T \mathbf{A}\mathbf{p}_j = 0.$$

To znači da je kod gradijentne metode vektor smjera u koraku $j+1$ ortogonalan s vektorom smjera u koraku j.

Primjer 1. *Zadana je funkcija* $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T \mathbf{A}\mathbf{x} - \mathbf{b}^T \mathbf{x}$, *gdje su*

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Jednostavno se vidi da je $\mathbf{x}^* = \left[\frac{1}{3}, \frac{1}{3}\right]^T$ *rješenje sustava linearnih jednadžbi. Pretpostavimo da gradijentnom metodom želimo odrediti točku globalnog minimuma funkcije* f. *U tu svrhu za početnu aproksimaciju uzimamo* $\mathbf{x}_0 = [5, -2]^T$. *Gradijentnu metodu izvršavamo sve do je* $\|\mathbf{A}\mathbf{x}_j - \mathbf{b}\|_2 \geq \varepsilon$, $j = 0, 1, 2, \ldots$. *Ako je* $\varepsilon = 10^{-4}$ *gradijentnom metodom za 17 koraka dobivamo* $\mathbf{x}_{17} = [0.333351, 0.333298]^T$. *Na Slici 1 prikazani su odgovarajući vektori smjerova gradijentne metode.*

2 Metoda konjugiranih gradijenata

Prirodna ideja je definirati vektore smjerova najbržeg silaska i to tako da metoda završi u što manje koraka. Pokazat ćemo da metoda konjugiranih gradijenata ima upravo to svojstvo te da završi u najviše n koraka.

Za simetričnu pozitivno definitnu matricu $\mathbf{A}$ definiramo preslikavanje $s : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, formulom

$$s(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{A} \mathbf{y}. \tag{3}$$

Može se pokazati da je preslikavanje s skalarni produkt kojeg označavamo s $s(\mathbf{x}, \mathbf{y}) =: \langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{A}}$ (vidi Domaću zadaću) i zovemo $\mathbf{A}-$skalarni produkt. Pri tome odgovarajuća normu definiramo s $\|\mathbf{x}\|_{\mathbf{A}} := \sqrt{\mathbf{x}^T \mathbf{A} \mathbf{x}}$.

Definicija 1. *Kažemo da su vektori $\mathbf{p}_0, \ldots, \mathbf{p}_{k-1}$ $\mathbf{A}-$konjugirani ili $\mathbf{A}-$ortogonalni, ako vrijedi da je $\mathbf{p}_j^T \mathbf{A} \mathbf{p}_i$ za svaki $i, j = 1, \ldots, k-1$ te $i \neq j$.*

Uočimo da je zbog simetričnosti dovoljno zahtijevati $\mathbf{p}_j^T \mathbf{A} \mathbf{p}_i = 0$ za svaki $0 \leq i < j \leq k-1$. Može se pokazati da je svaki skup $\{\mathbf{p}_j : j = 0, \ldots, k-1\}$, $\mathbf{A}-$konjugiranih vektora linearno nezavisan (vidi Domaću zadaću) te stoga možemo definirati potprostor

$$W_k = \begin{cases} \operatorname{span}\{\mathbf{p}_0, \ldots, \mathbf{p}_{k-1}\}, & k \geq 1 \\ \{\mathbf{0}\}, & k = 0 \end{cases},$$

te skup

$$U_k := \begin{cases} \mathbf{x}_0 + W_k = \{\mathbf{z} \in \mathbb{R}^n : \mathbf{z} = \mathbf{x}_0 + \mathbf{w}_k \in W_k\}, & k \geq 1 \\ \{\mathbf{x}_0\}, & k = 0. \end{cases}$$

Propozicija 1. *Neka je i fiksni prirodni broj te neka su $\{\mathbf{p_0}, \ldots, \mathbf{p}_{i-1}, \mathbf{p}_i\}$ vektori takvi da je*

$$\mathbf{p}_i^T \mathbf{A} \mathbf{p}_j = 0, \tag{4}$$

za svaki $j = 0, \ldots, i-1$. Prepostavimo da je $\mathbf{x}_0 \in \mathbb{R}^n$ te $\mathbf{x}_j \in \mathbb{R}^n$, $j = 1, \ldots, i-1$ niz vektora definiran rekurzivno s

$$\mathbf{x}_{j+1} = \mathbf{x}_j + \alpha_j \mathbf{p}_j,$$

$$\alpha_j = \operatorname{argmin}_{\alpha \in \mathbb{R}} f(\mathbf{x}_j + \alpha \mathbf{p}_j) = -\frac{\mathbf{p}_j^T (\mathbf{A}\mathbf{x}_j - \mathbf{b})}{\mathbf{p}_j^T \mathbf{A} \mathbf{p}_j} = -\frac{\mathbf{p}_j^T \mathbf{r}_j}{\mathbf{p}_j^T \mathbf{A} \mathbf{p}_j},$$

gdje su $\mathbf{r}_j := \mathbf{A}\mathbf{x}_j - \mathbf{b}$, $j = 0, \ldots, i-1$. Tada vrijedi

$$\mathbf{p}_i^T \mathbf{r}_i = \mathbf{p}_i^T (\mathbf{A}\mathbf{y} - \mathbf{b}), \ \text{za svaki } \mathbf{y} \in U_i.$$

Dokaz. Uočimo da je

$$\mathbf{x}_i = \mathbf{x}_{i-1} + \alpha_{i-1}\mathbf{p}_{i-1} = \mathbf{x}_{i-2} + \alpha_{i-2}\mathbf{p}_{i-2} + \alpha_{i-1}\mathbf{p}_{i-1} = \ldots = \mathbf{x}_0 + \sum_{j=1}^{i-1} \alpha_j \mathbf{p}_j \in U_i.$$

Za proizvoljni $\mathbf{y} \in U_i$ vrijedi da je $\mathbf{x}_i - \mathbf{y} \in W_i$ te stoga zbog svojstva (4) vrijedi $\mathbf{p}_i^T \mathbf{A}(\mathbf{x}_i - \mathbf{y}) = 0$. Konačno slijedi

$$\mathbf{p}_i^T (\mathbf{r}_i - \mathbf{A}\mathbf{y} + \mathbf{b}) = \mathbf{p}_i^T (\mathbf{A}\mathbf{x}_i - \mathbf{b} - \mathbf{A}\mathbf{y} + \mathbf{b}) = \mathbf{p}_i^T \mathbf{A}(\mathbf{x}_i - \mathbf{y}) = 0,$$

odakle slijedi tvrdnja.

U sljedeće teoremu pokazat ćemo da ako za vektore smjerova odaberemo $\mathbf{A}-$konjugirane vektore, onda jednodimenzionalna minimizacija duž vektora $\mathbf{p}_{k-1}$ (odnosno traženje vektora $\mathbf{x}_k$) odgovara minimizaciji duž cijelog skupa U_k.

Teorem 1. *Neka je* $\mathbf{x}_0 \in \mathbb{R}^n$ *te* $\mathbf{x}_{j+1} = \mathbf{x}_j + \alpha_j p_j$, $j = 0, \ldots, k-1$, *pri čemu su vektori smjerova* $\mathbf{p}_j$, $j = 0, \ldots, k-1$ $\mathbf{A}-$*konjugirani te* $\alpha_j \in \mathbb{R}$ *odgovarajuće duljine koraka. Onda je*

$$\mathbf{x}_j = \operatorname{argmin}_{x \in U_j} f(\mathbf{x}),\ j = 1, \ldots, k.$$

Dokaz. Tvrdnju dokazujemo metodom matematičke indukcije po k. Ako je $k = 1$, onda je $\mathbf{x}_1 = \mathbf{x}_0 + \alpha_0 \mathbf{p}_0$, gdje je $\alpha_0 = -\frac{\mathbf{p}_0^T \mathbf{r}_0}{\mathbf{p}_0^T \mathbf{A} \mathbf{p}_0}$, točka globalnog minimuma funkcije f na skupu $U_1 = \{\mathbf{x}_0 + \mathbf{w}_1 : \mathbf{w}_1 \in W_1 = \operatorname{span}\{\mathbf{p}_0\}\}$. Pretopstavimo da tvrdnja vrijedi za prirodni broj $k = i$, tj. da je

$$\mathbf{x}_j = \operatorname{argmin}_{\mathbf{y} \in U_j} f(\mathbf{y}),$$

za svaki $j = 1, \ldots, i$. Dokažimo da tvrdnja vijedi za $k = i + 1$ tj. da iz $\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha_i \mathbf{p}_i$ slijedi da je

$$\mathbf{x}_{i+1} = \operatorname{argmin}_{\mathbf{x} \in U_{i+1}} f(\mathbf{x}).$$

Iz definicije skupa U_{i+1} slijedi da se svaki $\mathbf{x} \in U_{i+1}$ može zapisati u obliku $\mathbf{x} = \mathbf{y} + \alpha \mathbf{p}_i$, pri čemu je $\alpha \in \mathbb{R}$ te $\mathbf{y} \in U_i$ te je stoga

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{y} + \alpha \mathbf{p}_i) = f(\mathbf{y}) + \alpha \mathbf{p}_i^T (\nabla f(\mathbf{y})) + \frac{1}{2}\alpha^2 \mathbf{p}_i^T \mathbf{A} \mathbf{p}_i \\ &= f(\mathbf{y}) + \alpha \mathbf{p}_i^T (\mathbf{A}\mathbf{y} - \mathbf{b}) + \frac{1}{2}\alpha^2 \mathbf{p}_i^T \mathbf{A} \mathbf{p}_i, \end{aligned}$$

odakle prema Propoziciji 1 slijedi da je

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

Konačno imamo

$$\min_{\mathbf{x} \in U_{i+1}} f(\mathbf{x}) = \min_{\mathbf{y} \in U_i} f(\mathbf{y}) + \min_{\alpha \in \mathbb{R}} \left(\alpha \mathbf{p}_i^T (\mathbf{A}\mathbf{x}_i - \mathbf{b}) + \frac{1}{2}\alpha^2 \mathbf{p}_i^T \mathbf{A} \mathbf{p}_i \right).$$

Prema pretpostavci indukcije slijedi da je $\operatorname{argmin}_{\mathbf{y} \in U_i} f(\mathbf{y}) = \mathbf{x}_i$. Također slijedi da je

$$\operatorname{argmin}_{\alpha \in \mathbb{R}} \left(\alpha \mathbf{p}_i^T (\mathbf{A}\mathbf{x}_i - \mathbf{b}) + \frac{1}{2}\alpha^2 \mathbf{p}_i^T \mathbf{A} \mathbf{p}_i \right) = -\frac{\mathbf{p}_i^T \mathbf{r}_i}{\mathbf{p}_i^T \mathbf{A} \mathbf{p}_i} = \alpha_i,$$

odakle slijedi tvrdnja.

U sljedećoj Propoziciji pokazat ćemo kako iz vektora reziduala $\mathbf{r}_j$, $j = 1, \ldots, k$, konstruirati $A-$konjugirane smjerove.

Propozicija 2. *Neka je* $\mathbf{p}_0 = -\mathbf{r}_0$ *te*

$$\mathbf{p}_k = -\mathbf{r}_k + \sum_{j=0}^{k-1} \frac{\mathbf{p}_j^T \mathbf{A} \mathbf{r_k}}{\mathbf{p}_j^T \mathbf{A} \mathbf{p}_j} \mathbf{p}_j, \quad k = 1, 2, \ldots. \tag{5}$$

Onda je $\mathbf{p}_j^T \mathbf{A} \mathbf{p}_m = 0$ *za svaki* $0 \leq m < j \leq k$.

Dokaz. Tvrdnju dokazujemo metodom matematičke indukcije po k. Za $k = 1$ imamo

$$\mathbf{p}_1^T \mathbf{A}\mathbf{p}_0 \;=\; \left(-\mathbf{r}_1 + \frac{\mathbf{p}_0^T \mathbf{A}\mathbf{r}_1}{\mathbf{p}_0^T \mathbf{A}\mathbf{p}_0}\mathbf{p}_0\right)^T \mathbf{A}\mathbf{p}_0 = -\mathbf{r}_1^T \mathbf{A}\mathbf{p}_0 + \mathbf{p}_0^T \mathbf{A}\mathbf{p}_0 \frac{\mathbf{p}_0^T \mathbf{A}\mathbf{r}_1}{\mathbf{p}_0^T \mathbf{A}\mathbf{p}_0} = 0.$$

Pretpostavimo da tvrdna vrijedi za prirodni broj $k = i$, odnosno da su vektori $\mathbf{p}_0, \ldots, \mathbf{p}_i$ $A-$konjugirani. Dokažimo da tvrdnja vrijedi za za $k = i + 1$ tj. da je $\mathbf{p}_{i+1}^T \mathbf{A}\mathbf{p}_m = 0$ za svaki $m \leq i$. Neka je m prirodni broj koji je manji od i. Onda je

$$\begin{aligned}\mathbf{p}_{i+1}^T \mathbf{A}\mathbf{p}_m \;&=\; -\mathbf{r}_{i+1}^T \mathbf{A}\mathbf{p}_m + \sum_{j=0}^{i} \frac{\mathbf{p}_j^T \mathbf{A}\mathbf{r}_{i+1}}{\mathbf{p}_j^T \mathbf{A}\mathbf{p}_j}\mathbf{p}_j^T \mathbf{A}\mathbf{p}_m \\ &=\; -\mathbf{r}_{i+1}^T \mathbf{A}\mathbf{p}_m + \frac{\mathbf{p}_m^T \mathbf{A}\mathbf{r}_{i+1}}{\mathbf{p}_m^T \mathbf{A}\mathbf{p}_m}\mathbf{p}_m^T \mathbf{A}\mathbf{p}_m = 0.\end{aligned}$$

U sljedećoj propoziciji pokazat ćemo da se konjugirani smjerovi mogu računati jednostavnije.

Propozicija 3. *Pretpostavimo da su* $\mathbf{p}_0, \ldots, \mathbf{p}_k$ *smjerovi dobiveni rekurzivno s (5). Onda je*

a) $W_k = span\{\mathbf{r}_0, \ldots, \mathbf{r}_{k-1}\}$

b) $\mathbf{r}_m^T \mathbf{r}_j = 0$ *za svaki* $0 \leq j < m \leq k$

c) $\mathbf{p}_k^T \mathbf{r}_j = -\mathbf{r}_k^T \mathbf{r}_k$ *za svaki* $0 \leq j \leq k$

d) Za vektor smjera $\mathbf{p}_k$ *vrijedi*

$$\mathbf{p}_k = -\mathbf{r}_k + \beta_{k-1}\mathbf{p}_{k-1}, \quad \textit{gdje je } \beta_{k-1} = \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{r}_{k-1}^T \mathbf{r}_{k-1}}.$$

e) Vrijedi $W_k = span\{\mathbf{r}_0, \ldots, \mathbf{A}^{k-1}\mathbf{r}_0\}$

Dokaz. Dokazat ćemo tvrdnju d). U tu svrhu vektor $\mathbf{p}_k \in W_{k+1}$ zapišimo kao linearnu kombinaciju vektora $\mathbf{r}_0, \ldots, \mathbf{r}_k$. Sukladno svojstvu b) vektori $\mathbf{r}_0, \ldots, \mathbf{r}_k$ čine ortogonalnu bazu od W_{k+1}. Vrijedi

$$\mathbf{p}_k = \sum_{j=0}^{k} \alpha_j \mathbf{r}_j,$$

pri čemu se iz uvjeta ortogonalnosti vektora $\mathbf{r}_0, \ldots, \mathbf{r}_k$ slijedi da je $\alpha_j = \frac{\mathbf{p}_k^T \mathbf{r}_j}{\mathbf{r}_j^T \mathbf{r}_j}$, $j = 0, \ldots, k$. Prema svojstvu c) vrijedi

$$\begin{aligned}\mathbf{p}_k \;&=\; -\sum_{j=0}^{k} \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{r}_j^T \mathbf{r}_j}\mathbf{r}_j = -\mathbf{r}_k - \sum_{j=0}^{k-1} \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{r}_j^T \mathbf{r}_j}\mathbf{r}_j = -\mathbf{r}_k - \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{r}_{k-1}^T \mathbf{r}_{k-1}} \sum_{j=0}^{k-1} \frac{\mathbf{r}_{k-1}^T \mathbf{r}_{k-1}}{\mathbf{r}_j^T \mathbf{r}_j}\mathbf{r}_j \\ &=\; -\mathbf{r}_k + \beta_{k-1} \sum_{j=0}^{k-1} \frac{\mathbf{p}_{k-1}^T \mathbf{r}_j}{\mathbf{r}_j^T \mathbf{r}_j}\mathbf{r}_j = -\mathbf{r}_k + \beta_{k-1}\mathbf{p}_{k-1}. \qquad (6)\end{aligned}$$

Posljednja jednakost u (6) iz raspisa vektora $\mathbf{p}_{k-1}$ u ortogonalnoj bazi $\{\mathbf{r}_0, \ldots, \mathbf{r}_{k-1}\}$.

Na osnovi prethodnih tvrdnji proizlazi sljedeći algoritam.

ALGORITAM - KONJUGIRANI GRADIJENTI

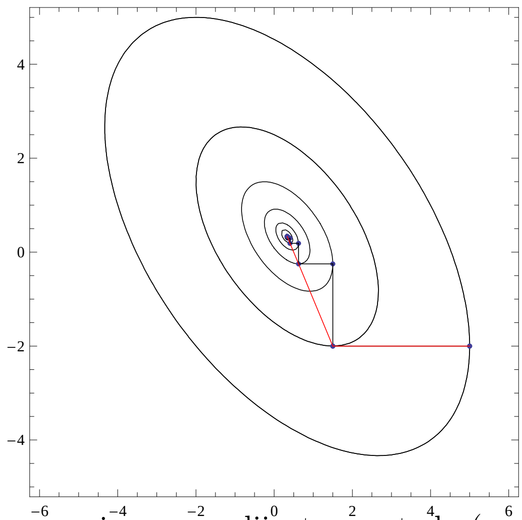
- $\mathbf{r}_0 = \mathbf{A}\mathbf{x}_0 - \mathbf{b}, \quad \mathbf{p}_0 = -\mathbf{r}_0,\ k = 0$
- While $\mathbf{r}_k \neq \mathbf{0}$ do

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

- end While

Primjedba 1. Ako je $\mathbf{A} \in \mathbb{R}^{n\times n}$, onda će zbog ortogonalnosti vektora reziduala iterativni postupak završiti u konačno mnogo (najviše n koraka).

Primjer 2. *Na Slici 2 prikazana je usporedba iterativnih procesa gradijentnom metodom te metodom konjugiranih gradijenata na minimizaciji funkcije iz Primjera 1.*



Slika 2: Usporedba vektora smjerova gradijentne metode (crno) i metode konjugiranih gradijenata (crveno)

2.1 Brzina konvergencije

Može se pokazat da brzina konvegrencije metode konjugiranih gradijenata ovisi o broju uvjetovanosti matrice $\mathbf{A}$, preciznije da vrijedi sljedeća tvrdnja.

Teorem 2. *(vidi Quarterioni at al (2000)) Ako je $\mathbf{x}^*$ egzaktno rješenje sustava linearnih jednadžbi $\mathbf{Ax} = \mathbf{b}$, gdje je $\mathbf{A} \in \mathbb{R}^{n \times n}$ pozitivno definitna matrica te $\mathbf{b} \in \mathbb{R}^n$ te $\mathbf{x}_k$ aproksimacija rješenja dobivena u $k-$toj iteraciji gradijentne metode. Vrijedi sljedeća ocjena*

$$\|\mathbf{x}^* - \mathbf{x}_k\|_{\mathbf{A}} \leq \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}\right) \|\mathbf{x}^* - \mathbf{x}_0\|_{\mathbf{A}},$$

pri čemu je κ broj uvjetovanosti matrice $\mathbf{A}$.

Teorem 3. *(vidi Luenberger (1993)) Ako je $\mathbf{x}^*$ egzaktno rješenje sustava linearnih jednadžbi $\mathbf{Ax} = \mathbf{b}$, gdje je $\mathbf{A} \in \mathbb{R}^{n \times n}$ pozitivno definitna matrica te $\mathbf{b} \in \mathbb{R}^n$ te $\mathbf{x}_k$ aproksimacija rješenja dobivena u $k-$toj iteraciji metode konjugiranih gradijenata. Vrijedi sljedeća ocjena*

$$\|\mathbf{x}^* - \mathbf{x}_k\|_{\mathbf{A}} \leq 2\left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}\right)^k \|\mathbf{x}^* - \mathbf{x}_0\|_{\mathbf{A}},$$

pri čemu je κ broj uvjetovanosti matrice $\mathbf{A}$.

Teorem 4. *(vidi Quarterioni at al (2000)) Ako je $\mathbf{x}^*$ egzaktno rješenje sustava linearnih jednadžbi $\mathbf{Ax} = \mathbf{b}$, gdje je $\mathbf{A} \in \mathbb{R}^{n \times n}$ pozitivno definitna matrica te $\mathbf{b} \in \mathbb{R}^n$ te $\mathbf{x}_k$ aproksimacija rješenja dobivena u $k-$toj iteraciji metode konjugiranih gradijenata. Vrijedi sljedeća ocjena*

$$\|\mathbf{x}^* - \mathbf{x}_k\|_{\mathbf{A}} \leq 2\frac{c^k}{1 + c^{2k}} \|\mathbf{x}^* - \mathbf{x}_0\|_{\mathbf{A}},$$

pri čemu je $c = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$ te κ broj uvjetovanosti matrice $\mathbf{A}$.

3 Proširenje na nelinearne funkcije

Metodu konjugiranih gradijenata moguće je proširiti i na nelinearne funkcije (vidi primjerice Quarterioni at al (2000)). U nastavku navodimo nekoliko modifikacija metode koje su prilagođene za nelinearne funkcije.

ALGORITAM FLETCHER-REEVES METHOD

1 $\mathbf{r}_0 = \nabla f(\mathbf{x}_0)$, $\mathbf{p}_0 = -\mathbf{r}_0$

2 For $k = 0, 1, \ldots, n-1$ do

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k, \text{ gdje je } \alpha_k = \text{argmin}_{\alpha \in \mathbb{R}} f(\mathbf{x}_k + \alpha \mathbf{p}_k)$$

$$\mathbf{r}_{k+1} = \nabla f(\mathbf{x}_{k+1})$$

$$\mathbf{p}_{k+1} = -\mathbf{r}_k + \beta_k \mathbf{p}_k, \text{ gdje je } \beta_k = \frac{\mathbf{r}_{k+1}^T \mathbf{r}_{k+1}}{\mathbf{r}_k^T \mathbf{r}_k}$$

3 $\mathbf{x}_0 = \mathbf{x}_n$ i skoči na korak 1.

Algoritam Polak-Ribiere metoda
Analogno kao Fletcher-Reeves method, razlika je u β_k:

$$\beta_k = \frac{(\mathbf{r}_{k+1} - \mathbf{r}_k)^T \mathbf{r}_{k+1}}{\mathbf{r}_k^T \mathbf{r}_k}.$$

Najčešći izbor u $\beta_k = \max\{0, \frac{(\mathbf{r}_{k+1} - \mathbf{r}_k)^T \mathbf{r}_{k+1}}{\mathbf{r}_k^T \mathbf{r}_k}\}$.
Hestenes-Stiefel metoda Analogno kao Fletcher-Reeves method, razlika je u β_k:

$$\beta_k = \frac{(\mathbf{r}_{k+1} - \mathbf{r}_k)^T \mathbf{r}_{k+1}}{\mathbf{p}_k^T (\mathbf{r}_{k+1} - \mathbf{r}_k)}$$

4 Domaća zadaća

Zadatak 1. *Dokažite da je svaka pozitivno definitna matrica* $\mathbf{M} \in \mathbb{R}^{n \times n}$ *regularna.*

Zadatak 2. *Neka je* $\mathbf{A} \in \mathbb{R}^{n \times n}$ *pozitivno definitna i simetrična matrica. Dokažite da je preslikavanje* $s : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *zadano s* $s(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{A} \mathbf{y}$ *skalarni produkt na* $\mathbb{R}^n$*.*

Zadatak 3. *Neka su* $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $\mathbf{x}_0 = [0, 0]^T$ *i* $r = 3$*. Prikažite skup*

$$K_{\mathbf{A}}(\mathbf{x}_0, r) = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x} - \mathbf{x}_0\|_{\mathbf{A}} \le r\}.$$

Zadatak 4. *Neka je* $\mathbf{A} \in \mathbb{R}^{n \times n}$ *te neka je* $\{\mathbf{p}_j \in \mathbb{R}^n : j = 0, \ldots, k-1\}$ *skup* $\mathbf{A}-$*konjugiranih vektora. Dokažite da su vektori* $\mathbf{p}_j$, $j = 0, \ldots, k-1$ *linearno nezavisni.*

Zadatak 5. *Zadani su pozitivno definitna matrica* $\mathbf{A} \in \mathbb{R}^{n \times n}$*, vektori* $\mathbf{b} \in \mathbb{R}^n$ *te* $\mathbf{x}_0 \in \mathbb{R}^n$*. Definiramo* $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$, $\alpha_k = -\frac{\mathbf{r}_k^T \mathbf{p}_k}{\mathbf{p}_k^T \mathbf{A} \mathbf{p}_k}$, $\mathbf{r}_k = \mathbf{A}\mathbf{x}_k - \mathbf{b}$, $k = 1, 2, \ldots$*. Pretpostavimo da su* $\mathbf{p}_0, \ldots, \mathbf{p}_k$ *vektori dobiveni na sljedeći način*

$$\mathbf{p}_0 = -\mathbf{r}_0, \quad \mathbf{p}_k = -\mathbf{r}_k + \sum_{j=0}^{k-1} \frac{\mathbf{p}_j^T \mathbf{A} \mathbf{r}_\mathbf{k}}{\mathbf{p}_j^T \mathbf{A} \mathbf{p}_j} \mathbf{p}_j, \quad k = 1, 2, \ldots.$$

.

Dokažite da vrijedi

a) $W_k :=$ *span*$\{\mathbf{p}_0, \ldots, \mathbf{p}_{k-1} =$ *span*$\{\mathbf{r}_0, \ldots, \mathbf{r}_{k-1}\}$

b) $\mathbf{r}_m^T \mathbf{r}_j = 0$ *za svaki* $0 \le j < m \le k$

c) $\mathbf{p}_k^T \mathbf{r}_j = -\mathbf{r}_k^T \mathbf{r}_k$ *za svaki* $0 \le j \le k$

d) $W_k =$ *span*$\{\mathbf{r}_0, \ldots, \mathbf{A}^{k-1} \mathbf{r}_0\}$

Zadatak 6. *Prilagodite metodu konjugiranih gradijenata za minimizaciju nelinearne funkcije, tako da koristite ideju modificirane Wolfove metode jednodimenzionalne minimizacije prema članku* **?***. Metodu ilustrirajte na primjerima različitih nelinearnih funkcija.*

Literatura

A. Quarteroni, R. Sacco, F. Saleri, Numerical Mathematics, Springer-Verlag New York, 2000

D. G. Luenberger, Intro. to Linear & Nonlinear Programming, Addison-Wesley, 1993

Gaohang Yu, Lutai Guan, Zengxin Wei, *Globally convergent Polak–Ribičre–Polyak conjugate gradient methods under a modified Wolfe line search*, Applied Mathematics and Computation, **215** (2009), 3082–3090