

# Metoda gradientów sprzężonych — dowód

PFG

listopad 2025

Niech macierz  $\mathbf{A} \in \mathbb{R}^{N \times N}$  będzie symetryczna i dodatnio określona. Niech  $\mathbf{r}_1 \in \mathbb{R}^N$  będzie dowolnym wektorem takim, że  $\|\mathbf{r}_1\| \neq 0$  i niech  $\mathbf{p}_1 = \mathbf{r}_1$ . Definiujemy następującą iterację:

$$\alpha_k = \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{p}_k^T \mathbf{A} \mathbf{p}_k}, \quad (1a)$$

$$\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{A} \mathbf{p}_k, \quad (1b)$$

$$\beta_k = \frac{\mathbf{r}_{k+1}^T \mathbf{r}_{k+1}}{\mathbf{r}_k^T \mathbf{r}_k}, \quad (1c)$$

$$\mathbf{p}_{k+1} = \mathbf{r}_{k+1} + \beta_k \mathbf{p}_k. \quad (1d)$$

Udowodnimy indukcyjnie, że dla każdego  $i, j, i > j$ , zachodzi

$$\mathbf{r}_i^T \mathbf{r}_j = 0, \quad (2a)$$

$$\mathbf{r}_i^T \mathbf{p}_j = 0, \quad (2b)$$

$$\mathbf{p}_i^T \mathbf{A} \mathbf{p}_j = 0. \quad (2c)$$

Zwróćmy uwagę, że wszystkie wielkości stojące po lewej stronie (2) są liczbami (skalarami).

## 1 Pierwszy krok

W pierwszym kroku dowodu indukcyjnego wygenerujemy wektory  $\mathbf{r}_2, \mathbf{p}_2$  i pokażemy, że spełniają one relacje (2) z wektorami  $\mathbf{r}_1, \mathbf{p}_1$ .

A. Obliczmy

$$\mathbf{r}_2^T \mathbf{r}_1 = (\mathbf{r}_1 - \alpha_1 \mathbf{A} \mathbf{p}_1)^T \mathbf{r}_1 = (\mathbf{r}_1^T - \alpha_1 \mathbf{p}_1^T \mathbf{A}^T) \mathbf{r}_1 = (\mathbf{r}_1^T - \alpha_1 \mathbf{r}_1^T \mathbf{A}) \mathbf{r}_1 \quad (3a)$$

W ostatniej równości skorzystaliśmy z faktu, że  $\mathbf{p}_1 = \mathbf{r}_1$  oraz że macierz  $\mathbf{A}$  jest symetryczna; na dalszych etapach dowodu będziemy ten ostatni krok wykonywać bez komentarza. Wykonajmy teraz mnożenie przez wyrażenie w nawiasie i skorzystajmy z postaci  $\alpha_1$ . Otrzymujemy

$$\mathbf{r}_2^T \mathbf{r}_1 = \mathbf{r}_1^T \mathbf{r}_1 - \frac{\mathbf{r}_1^T \mathbf{r}_1}{\mathbf{p}_1^T \mathbf{A} \mathbf{p}_1} \mathbf{r}_1^T \mathbf{A} \mathbf{r}_1 = \mathbf{r}_1^T \mathbf{r}_1 - \frac{\mathbf{r}_1^T \mathbf{r}_1}{\mathbf{r}_1^T \mathbf{A} \mathbf{r}_1} \mathbf{r}_1^T \mathbf{A} \mathbf{r}_1 \quad (3b)$$

gdzie znów skorzystaliśmy z tego, że  $\mathbf{p}_1 = \mathbf{r}_1$ . W tej sytuacji mianownik ułamka upraszcza się z wielkością stojącą za ułamkiem i ostatecznie

$$\mathbf{r}_2^T \mathbf{r}_1 = 0. \quad (3c)$$

B. Ten sam rachunek pokazuje, że  $\mathbf{r}_2^T \mathbf{p}_1 = 0$ , gdyż  $\mathbf{p}_1 = \mathbf{r}_1$ .

C. Teraz obliczmy

$$\mathbf{p}_2^T \mathbf{A} \mathbf{p}_1 = (\mathbf{r}_2 + \beta_1 \mathbf{p}_1)^T \mathbf{A} \mathbf{p}_1 = (\mathbf{r}_2 + \beta_1 \mathbf{r}_1)^T \mathbf{A} \mathbf{p}_1, \quad (4a)$$

gdzie ponownie skorzystaliśmy z faktu, że  $\mathbf{p}_1 = \mathbf{r}_1$ . Wyrażenie  $\mathbf{A} \mathbf{p}_1$  wyliczamy z (1b):

$$\mathbf{A} \mathbf{p}_1 = -\frac{1}{\alpha_1} (\mathbf{r}_2 - \mathbf{r}_1). \quad (4b)$$

Wobec tego

$$\mathbf{p}_2^T \mathbf{A} \mathbf{p}_1 = -\frac{1}{\alpha_1} (\mathbf{r}_2 + \beta_1 \mathbf{r}_1) (\mathbf{r}_2 - \mathbf{r}_1)^T (\mathbf{r}_2 - \mathbf{r}_1) = -\frac{1}{\alpha_1} (\mathbf{r}_2^T \mathbf{r}_2 - \mathbf{r}_2^T \mathbf{r}_1 + \beta_1 \mathbf{r}_1^T \mathbf{r}_2 - \beta_1 \mathbf{r}_1^T \mathbf{r}_1) \quad (4c)$$

W (3c) już pokazaliśmy, że  $\mathbf{r}_2^T \mathbf{r}_1 = 0$ , a więc drugi i trzeci wyraz znikają. Skorzystajmy z jawnej postaci  $\beta_1$ .

$$\mathbf{p}_2^T \mathbf{A} \mathbf{p}_1 = -\frac{1}{\alpha_1} \left( \mathbf{r}_2^T \mathbf{r}_2 - \frac{\mathbf{r}_2^T \mathbf{r}_2}{\mathbf{r}_1^T \mathbf{r}_1} \mathbf{r}_1^T \mathbf{r}_1 \right) = 0, \quad (4d)$$

gdź wyrażenie w nawiasie redukuje się.

W ten sposób wykonaliśmy pierwszy krok dowodu indukcyjnego: wykazaliśmy, że nowo wygenerowane wektory  $\mathbf{r}_2, \mathbf{p}_2$  spełniają zależności (2) z wektorami  $\mathbf{r}_1, \mathbf{p}_1$ .

## 2 Założenie indukcyjne

Zakładamy, że zgodnie z (1) wygenerowaliśmy wektory  $\mathbf{r}_k, \mathbf{p}_k$  i młodsze i że **spełniają one relacje (2) aż do pokolenia  $k$  włącznie**.

Wygenerujemy teraz wektory  $\mathbf{r}_{k+1}, \mathbf{p}_{k+1}$  i sprawdzimy, że przy tym założeniu spełniają one relacje (2) ze swoimi poprzednikami z młodszych pokoleń.

## 3 Krok indukcyjny

D. Obliczmy

$$\mathbf{r}_{k+1}^T \mathbf{r}_k = (\mathbf{r}_k - \alpha_k \mathbf{A} \mathbf{p}_k)^T \mathbf{r}_k = \mathbf{r}_k^T \mathbf{r}_k - \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{p}_k^T \mathbf{A} \mathbf{p}_k} \mathbf{p}_k^T \mathbf{A} \mathbf{r}_k. \quad (5a)$$

Gdyby mianownik dało się uprościć z wyrażeniem stojącym za ułamkiem, prawa strona zredukowałaby się do zera. Ale czy wskazane wyrażenia są równe? Z (1d) ze wskaźnikiem obniżonym o 1 wyliczamy

$$\mathbf{r}_k = \mathbf{p}_k - \beta_{k-1} \mathbf{p}_{k-1}, \quad (5b)$$

a zatem

$$\mathbf{p}_k^T \mathbf{A} \mathbf{r}_k = \mathbf{p}_k^T \mathbf{A} (\mathbf{p}_k - \beta_{k-1} \mathbf{p}_{k-1}) = \mathbf{p}_k^T \mathbf{A} \mathbf{p}_k - \beta_{k-1} \mathbf{p}_k^T \mathbf{A} \mathbf{p}_{k-1}. \quad (5c)$$

Ostatnie wyrażenie znika na mocy założenia indukcyjnego (2c), czyli

$$\mathbf{p}_k^T \mathbf{A} \mathbf{r}_k = \mathbf{p}_k^T \mathbf{A} \mathbf{p}_k, \quad (5d)$$

a więc

$$\mathbf{r}_{k+1}^T \mathbf{r}_k = \mathbf{r}_k^T \mathbf{r}_k - \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{p}_k^T \mathbf{A} \mathbf{p}_k} \mathbf{p}_k^T \mathbf{A} \mathbf{p}_k = 0. \quad (5e)$$

Uwaga! Wyrażenie (5d) **nie** oznacza, że wektory  $\mathbf{p}_k, \mathbf{r}_k$  są równe! Wektory te mogą się różnić o składniki ortogonalne do wektora  $\mathbf{A} \mathbf{p}_k$ .

E. W poprzednim punkcie pokazaliśmy, że wektor  $\mathbf{r}_{k+1}$  jest ortogonalny do swojego bezpośredniego poprzednika. A co z młodszymi wektorami? Weźmy  $l < k$  i obliczmy

$$\mathbf{r}_{k+1}^T \mathbf{r}_l = (\mathbf{r}_k - \alpha_k \mathbf{A} \mathbf{p}_k)^T \mathbf{r}_l = \mathbf{r}_k^T \mathbf{r}_l - \alpha_k \mathbf{p}_k^T \mathbf{A} \mathbf{r}_l. \quad (6a)$$

Pierwszy wyraz po prawej znika na mocy założenia indukcyjnego (2a). Pozostały wyraz przekształcamy korzystając z (1d) z odpowiednio obniżonymi wskaźnikami:

$$\mathbf{r}_{k+1}^T \mathbf{r}_l = -\alpha_k \mathbf{p}_k^T \mathbf{A} \mathbf{r}_l = -\alpha_k \mathbf{p}_k^T \mathbf{A} (\mathbf{p}_l - \beta_{l-1} \mathbf{p}_{l-1}) = -\alpha_k \mathbf{p}_k^T \mathbf{A} \mathbf{p}_l + \alpha_k \beta_{l-1} \mathbf{p}_k^T \mathbf{A} \mathbf{p}_{l-1} = 0, \quad (6b)$$

gdyż oba wyrażenia znikają na mocy założenia indukcyjnego (2c).

F. Obliczamy

$$\mathbf{r}_{k+1}^T \mathbf{p}_k = (\mathbf{r}_k^T - \alpha_k \mathbf{p}_k^T \mathbf{A}) \mathbf{p}_k = \mathbf{r}_k^T \mathbf{p}_k - \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{p}_k^T \mathbf{A} \mathbf{p}_k} \mathbf{p}_k^T \mathbf{A} \mathbf{p}_k = \mathbf{r}_k^T (\mathbf{p}_k - \mathbf{r}_k). \quad (7a)$$

Wyrażenie w nawiasie obliczamy na podstawie (1d) ze wskaźnikiem obniżonym o 1.

$$\mathbf{r}_{k+1}^T \mathbf{p}_k = \beta_{k-1} \mathbf{r}_k^T \mathbf{p}_{k-1} = 0, \quad (7b)$$

co wynika z założenia indukcyjnego (2b).

G. Niech  $l < k$ . Obliczamy

$$\mathbf{r}_{k+1}^T \mathbf{p}_l = \mathbf{r}_k^T \mathbf{p}_l - \alpha_k \mathbf{p}_k^T \mathbf{A} \mathbf{p}_l = 0. \quad (8)$$

Pierwszy wyraz po prawej stronie znika na mocy założenia indukcyjnego (2b), drugi na mocy założenia indukcyjnego (2c).

H. Wreszcie trzeba sprawdzić warunek sprzężenia.

$$\mathbf{p}_{k+1}^T \mathbf{A} \mathbf{p}_k = (\mathbf{r}_{k+1}^T + \beta_k \mathbf{p}_k^T) \mathbf{A} \mathbf{p}_k. \quad (9a)$$

Wyrażenie  $\mathbf{A} \mathbf{p}_k$  wyliczamy z (1b).

$$\begin{aligned} \mathbf{p}_{k+1}^T \mathbf{A} \mathbf{p}_k &= \frac{-1}{\alpha_k} (\mathbf{r}_{k+1}^T + \beta_k \mathbf{p}_k^T) (\mathbf{r}_{k+1} - \mathbf{r}_k) \\ &= \frac{-1}{\alpha_k} (\mathbf{r}_{k+1}^T \mathbf{r}_{k+1} - \mathbf{r}_{k+1}^T \mathbf{r}_k + \beta_k \mathbf{p}_k^T \mathbf{r}_{k+1} - \beta_k \mathbf{p}_k^T \mathbf{r}_k). \end{aligned} \quad (9b)$$

Pokazaliśmy już, że  $\mathbf{r}_{k+1}^T \mathbf{r}_k = 0$ , zobacz (5e), a także, że  $\mathbf{r}_{k+1}^T \mathbf{p}_k = 0$ , zobacz (7b), a zatem drugi i trzeci wyraz znikają. Pozostaje

$$\mathbf{p}_{k+1}^T \mathbf{A} \mathbf{p}_k = \frac{-1}{\alpha_k} \left( \mathbf{r}_{k+1}^T \mathbf{r}_{k+1} - \frac{\mathbf{r}_{k+1}^T \mathbf{r}_{k+1}}{\mathbf{r}_k^T \mathbf{r}_k} \mathbf{p}_k^T \mathbf{r}_k \right), \quad (9c)$$

gdzie skorzystaliśmy z jawnej postaci  $\beta_k$ . Jeśli chodzi o ostatni człon, to

$$\mathbf{p}_k^T \mathbf{r}_k = (\mathbf{r}_k^T + \beta_{k-1} \mathbf{p}_{k-1}^T) \mathbf{r}_k = \mathbf{r}_k^T \mathbf{r}_k, \quad (9d)$$

gdyż  $\mathbf{r}_k^T \mathbf{p}_{k-1} = 0$  na mocy założenia indukcyjnego (2b). Zatem

$$\mathbf{p}_{k+1}^T \mathbf{A} \mathbf{p}_k = \frac{-1}{\alpha_k} \left( \mathbf{r}_{k+1}^T \mathbf{r}_{k+1} - \frac{\mathbf{r}_{k+1}^T \mathbf{r}_{k+1}}{\mathbf{r}_k^T \mathbf{r}_k} \mathbf{r}_k^T \mathbf{r}_k \right) = 0. \quad (9e)$$

I. Na koniec bierzemy  $l < k$  i obliczamy

$$\mathbf{p}_{k+1}^T \mathbf{A} \mathbf{p}_l = (\mathbf{r}_{k+1}^T + \beta_k \mathbf{p}_k^T) \mathbf{A} \mathbf{p}_l = \mathbf{r}_{k+1}^T \mathbf{A} \mathbf{p}_l + \beta_k \mathbf{p}_k^T \mathbf{A} \mathbf{p}_l. \quad (10a)$$

Ostatnie wyrażenie znika na mocy założenia indukcyjnego (2c).  $\mathbf{A} \mathbf{p}_l$  wyliczamy z (1b) z odpowiednio obniżonymi wskaźnikami.

$$\mathbf{p}_{k+1}^T \mathbf{A} \mathbf{p}_l = \frac{-1}{\alpha_l} \mathbf{r}_{k+1}^T (\mathbf{r}_{l+1} - \mathbf{r}_l) = 0, \quad (10b)$$

gdyż  $l < k$ ,  $l + 1 \leq k$ , a już pokazaliśmy, że  $\mathbf{r}_{k+1}$  jest ortogonalny do wszystkich poprzedzających go wektorów  $\mathbf{r}_j$ .

Ten punkt kończy dowód ☺.

## 4 Zakończenie iteracji

Iteracje (1) prowadzimy tak długo, aż  $\|\mathbf{r}_k\| \neq 0$ . Udowodniona już własność (2a) stanowi, kolejne wektory  $\mathbf{r}_i$  są wzajemnie ortogonalne. Tymczasem w przestrzeni  $N$ -wymiarowej nie da się zdefiniować więcej, niż  $N$  wzajemnie prostopadłych kierunków. Wektor  $\mathbf{r}_{N+1}$  nie może być niezależny od swoich poprzedników, a zarazem musi być do nich ortogonalny. Jedynym wektorem, który to spełnia, jest wektor zerowy: własność (2a) zapewnia, że  $\mathbf{r}_{N+1} \equiv 0$ .

## 5 Symetria i dodatnia określoność?

Z symetrii macierzy  $\mathbf{A}$ ,  $\mathbf{A}^T = \mathbf{A}$ , korzystaliśmy wielokrotnie w trakcie dowodu, po raz pierwszy w kroku (3a). A gdzie wykorzystywana jest dodatnia określoność macierzy? Otóż do samego zdefiniowania iteracji (1). Dodatnia określoność macierzy  $\mathbf{A}$  gwarantuje, że w wyrażeniu (1a) nie dzielimy przez zero. Bez tego wymagania cała iteracja mogłaby się załamać.

## 6 Zastosowanie do rozwiązywania równań liniowych

Przypuśćmy, że chcemy rozwiązać układ równań

$$\mathbf{A} \mathbf{x} = \mathbf{b}, \quad (11a)$$

przy czym macierz  $\mathbf{A}$  jest symetryczna i dodatnio określona. Weźmy *dowolny* wektor  $\mathbf{x}_1$  i zdefiniujmy  $\mathbf{r}_1 = \mathbf{b} - \mathbf{A} \mathbf{x}_1$ . Jeżeli  $\mathbf{r}_1 = 0$  nic więcej nie musimy robić:  $\mathbf{x}_1$  jest rozwiązaniem równania (11a). Jeśli  $\mathbf{r}_1 \neq 0$ , możemy go użyć do zainicjowania iteracji (1). Definiujemy

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k. \quad (11b)$$

Twierdzimy, że  $\mathbf{x}_{N+1}$  jest rozwiązaniem równania (11a).

W tym celu wystarczy pokazać, że

$$\mathbf{r}_k = \mathbf{b} - \mathbf{A} \mathbf{x}_k. \quad (11c)$$

Jeżeli (11c) zachodzi, to ponieważ  $\mathbf{r}_{N+1} = 0$ ,  $\mathbf{x}_{N+1}$  jest poszukiwanym rozwiązaniem.

Dowód (11c) także prowadzimy indukcyjnie.

Pierwszy krok:

$$\mathbf{x}_2 = \mathbf{x}_1 + \alpha_1 \mathbf{p}_1, \quad (11d)$$

$$\begin{aligned} \mathbf{b} - \mathbf{A}\mathbf{x}_2 &= \mathbf{b} - \mathbf{A}(\mathbf{x}_1 + \alpha_1 \mathbf{p}_1) \\ &= \mathbf{b} - \mathbf{A}\mathbf{x}_1 - \alpha_1 \mathbf{A}\mathbf{p}_1 = \mathbf{r}_1 - \alpha_1 \mathbf{A}\mathbf{p}_1 = \mathbf{r}_2 \end{aligned} \quad (11e)$$

Krok indukcyjny: Załóżmy, że (11c) zachodzi aż do pokolenia  $k$  włącznie. Wówczas

$$\begin{aligned} \mathbf{b} - \mathbf{A}\mathbf{x}_{k+1} &= \mathbf{b} - \mathbf{A}(\mathbf{x}_k + \alpha_k \mathbf{p}_k) \\ &= \mathbf{b} - \mathbf{A}\mathbf{x}_k - \alpha_k \mathbf{A}\mathbf{p}_k = \mathbf{r}_k - \alpha_k \mathbf{A}\mathbf{p}_k = \mathbf{r}_{k+1}. \end{aligned} \quad (11f)$$