Metody iteracyjne

Na tych laboratoriach skupimy się na rozwiązaniu układu równań:

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

W tym celu rozpatrzymy kilka metod, których działanie sprawdzimy na układzie równań otrzymanym na poprzednich zajęciach dzięki metodzie MES.
Naszym celem będzie napisanie funkcji `Solve`, która zastąpi funkcję `Gauss` i wyznaczy wartość wektora $\mathbf{x}$. Nie będziemy jednak tego układu rozwiązywać metodą bezpośrednią, taką jak eliminacja Gaussa, ale **metodą iteracyjną**. Skonstruujemy ciąg $\left(\mathbf{x}^k\right)$, którego elementy $\mathbf{x}^k$ będą dążyły do rozwiązania dokładnego $\mathbf{x}$.

Niezbędne definicje

- Wektorem błędu w k-tym kroku nazywamy:

$$\mathbf{e}^k = \mathbf{x} - \mathbf{x}^k$$

- Wektorem residualnym (**residuum**) w k-tym kroku nazywamy:

$$\mathbf{r}^k = \mathbf{b} - \mathbf{b}^k$$

Łatwo zauważyć, że zachodzi zależność $\mathbf{r}^k = \mathbf{A}\mathbf{e}^k$.
Widać także, że skoro $\mathbf{x}^k$ dąży do $\mathbf{x}$ to $\mathbf{r}^k$ dąży do zera.

Zadanie

Wyznacz residuum dla zadania z poprzednich zajęć. Następnie oblicz i wyświetl jego normę: $\|\mathbf{r}\| = \sqrt{\mathbf{r}^T\mathbf{r}}$ (napisz funkcję liczącą normę wektora `norm(int, double *)`). Ile wynosi ta norma przed i po rozwiązaniu układu metodą eliminacji Gaussa?

Początki

Weźmy dowolny wektor $\mathbf{x}^0$ i obliczmy odpowiadający mu wektor prawych stron $\mathbf{b}^0 = \mathbf{A}\mathbf{x}^0$. Różnica między „prawdziwym"'wektorem $\mathbf{b}$ a przybliżeniem jest wtedy równa

$$\mathbf{r}^0 = \mathbf{b} - \mathbf{b}^0 = \mathbf{A}\mathbf{x} - \mathbf{A}\mathbf{x}^0 = \mathbf{A}\mathbf{e}^0$$

Zatem różnica między „prawdziwym"'rozwiązaniem a przybliżonym $\mathbf{e}^0 = \mathbf{A}^{-1}\mathbf{r}^0$. Co ostatecznie pozwala nam zapisać: $\mathbf{x} = \mathbf{x}^0 + \mathbf{A}^{-1}\mathbf{r}^0$. Nie mamy jednak $\mathbf{A}^{-1}$ (w tym rzecz). Zamiast niej użyjemy macierzy $\mathbf{M}^{-1}$. Wtedy jednak nie dostaniemy dokładnej wartości[1] $\mathbf{x}$ a jedynie przybliżenie. Prowadzi to nas do wzoru:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{p}^k = \mathbf{x}^k + \mathbf{M}^{-1}\mathbf{r}^k$$

Wektor $\mathbf{p}^k$ jest „poprawką"'w k-tej iteracji a macierz $\mathbf{M}$ nazywamy **preconditioner'em**. Najlepiej byłoby gdyby macierz $\mathbf{M}$ była „podobna" do macierzy $\mathbf{A}$ a jednocześnie łatwo odwracalna.

Rozpatrzmy układ $\mathbf{r}^k = \mathbf{A}\mathbf{p}^k$. Widać, że jeśli pominiemy większość jego elementów:

$$\begin{cases} A_{11}p_1^k & + & A_{12}p_2^k & + & A_{13}p_3^k & + & \ldots & + & A_{1N}p_N^k & = & r_1^k \\ A_{21}p_1^k & + & A_{22}p_2^k & + & A_{23}p_3^k & + & \ldots & + & A_{2N}p_N^k & = & r_2^k \\ A_{31}p_1^k & + & A_{32}p_2^k & + & A_{33}p_3^k & + & \ldots & + & A_{3N}p_N^k & = & r_3^k \\ \vdots & & \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ A_{N1}p_1^k & + & A_{N2}p_2^k & + & A_{N3}p_3^k & + & \ldots & + & A_{NN}p_N^k & = & r_N^k \end{cases}$$

dostaniemy prosty wzór na $\mathbf{p}^k$:

$$p_i^k = \frac{1}{A_{ii}} r_i^k$$

Jest to równoważne z wzięciem za macierz $\mathbf{M}$ diagonalnej części macierzy $\mathbf{A}$. Ten prosty schemat iteracji z powyższą poprawką nazywamy **metodą Jacobiego**.

Zadania

1. Zaimplementuj metodę Jacobiego i wykonaj np. 1000 iteracji zaczynając od $\mathbf{x}^0 = 0$. W każdej iteracji wyświetl normę residuum. Napisz także funkcję `res_draw(double)`, która posłuży do wykonania wykresu zbieżności.
2. Taki proces iteracyjny nie zbiega się. Wprowadź współczynnik α, który „przytłumi" wykonywane iteracje:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha\mathbf{p}^k$$

3. Sprawdź zbieżność tego schematu dla różnych wartości α. Sprawdź liczby 0.5, 0.9, 1.1 i 2.
4. Wydziel z funkcji `Solve` część odpowiedzialną za mnożenie przez $\mathbf{A}$: `Mult(int N, double **A, double *x, double *r)` i preconditioner: `Precond(int N, double **A, double *x, double *p)`.

[1] I tak nie dostalibyśmy dokładnej wartości ze względu na błędy numeryczne.

Spróbujmy poprawić nasz schemat biorąc lepszy preconditioner. Zauważmy, że obliczając p_2^k mamy już obliczone p_1^k i możemy go użyć. Nie musimy zatem pomijać elementów układu „pod diagonalą":

$$\begin{cases} A_{11}p_1^k & + & A_{12}p_2^k & + & A_{13}p_3^k & + & \dots & + & A_{1N}p_N^k & = & r_1^k \\ A_{21}p_1^k & + & A_{22}p_2^k & + & A_{23}p_3^k & + & \dots & + & A_{2N}p_N^k & = & r_2^k \\ A_{31}p_1^k & + & A_{32}p_2^k & + & A_{33}p_3^k & + & \dots & + & A_{3N}p_N^k & = & r_3^k \\ \vdots & & \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ A_{N1}p_1^k & + & A_{N2}p_2^k & + & A_{N3}p_3^k & + & \dots & + & A_{NN}p_N^k & = & r_N^k \end{cases}$$

Daje nam to prosty wzór na $\mathbf{p}^k$:

$$p_i^k = \frac{1}{A_{ii}}\left(r_i - \sum_{j=1}^{i-1} A_{ij}p_j^k\right)$$

Gdy $\alpha = 1$ schemat taki nazywamy **metodą Gaussa-Seidla**.

Zadania

1. Wypróbuj nowy wzór na $\mathbf{p}$. Sprawdź różne wartości α.

Schematy, dla których $\alpha > 1$ nazywamy metodami **Successive Over-Relaxation** (SOR).

Dobieramy α

Widać wyraźnie, że zbieżność bardzo zależy od wartości współczynnika α i jasnym jest, że najlepiej byłoby dobierać ten współczynnik w każdej iteracji:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha^k \mathbf{p}^k$$

Zastanówmy się teraz jak będzie się zmieniało residuum w zależności od kroku. Jeśli pomnożymy powyższy wzór przez $-\mathbf{A}$ a następnie dodamy $\mathbf{b}$ i skorzystamy z definicji residuum otrzymamy:

$$\mathbf{r}^{k+1} = \mathbf{r}^k - \alpha^k \mathbf{A}\mathbf{p}^k$$

Kwadrat normy tego residuum jest równy:

$$\|\mathbf{r}^{k+1}\| = (\mathbf{r}^{k+1})^T\mathbf{r}^{k+1} = (\mathbf{r}^k - \alpha^k\mathbf{A}\mathbf{p}^k)^T(\mathbf{r}^k - \alpha^k\mathbf{A}\mathbf{p}^k) =$$

$$(\mathbf{r}^k)^T\mathbf{r}^k - 2\alpha^k(\mathbf{r}^k)^T\mathbf{A}\mathbf{p}^k + (\alpha^k)^2(\mathbf{A}\mathbf{p}^k)^T\mathbf{A}\mathbf{p}^k$$

Widać, że kwadrat normy jest kwadratową funkcją α^k a współczynnik przed $(\alpha^k)^2$ jest dodatni. Oznacza to, że funkcja ta ma minimum. Obliczamy pochodną po α^k:

$$\frac{d}{d\alpha^k}\left(\|\mathbf{r}^{k+1}\|\right) = -2(\mathbf{r}^k)^T\mathbf{A}\mathbf{p}^k + 2\alpha^k(\mathbf{A}\mathbf{p}^k)^T\mathbf{A}\mathbf{p}^k$$

i przyrównujemy do zera co ostatecznie daje wartość:

$$\alpha^k = \frac{(\mathbf{r}^k)^T\mathbf{A}\mathbf{p}^k}{(\mathbf{A}\mathbf{p}^k)^T\mathbf{A}\mathbf{p}^k}$$

Schemat z taką wartością α^k nazywamy **metodą najmniejszych residuów** (Minimal Residual Method — **MINRES**).

Zadania

1. Sprawdź zbieżność dla nowego α^k. W tym celu:
   - Wyznacz wektor $\mathbf{A}\mathbf{p}^k$.
   - Zauważ, że wyrażenie typu $\mathbf{a}^T\mathbf{b}$ to iloczyn skalarny dwóch wektorów $\mathbf{a}^T\mathbf{b} = \mathbf{a}\cdot\mathbf{b}$. Napisz funkcję `skal(int, double *, double *)` liczącą iloczyn skalarny i oblicz α^k z powyższego wzoru.

Wykorzystujemy historię

Do tej pory ignorowaliśmy informację o poprawkach z poprzednich iteracji i poprawkę w k-tym kroku obliczaliśmy ze wzoru

$$\mathbf{p}^k = \mathbf{M}^{-1}\mathbf{r}^k$$

Zmienimy to przez wykorzystanie informacji o poprawce z $k-1$ kroku:

$$\mathbf{p}^k = \mathbf{p}^k - \beta^k\mathbf{p}^{k-1}$$

Teraz wzór na nowe residuum będzie miał postać:

$$\mathbf{r}^{k+1} = \mathbf{r}^k - \alpha^k\mathbf{A}\left(\mathbf{p}^k - \beta^k\mathbf{p}^{k-1}\right)$$

Musimy jeszcze wyznaczyć wartość nowego współczynnika β^k.

Zadania

1. Wypisz wzór na $\|\mathbf{r}^{k+1}\|$ i zróżniczkuj go po współczynniku β^k. Przyjmij, że $(\mathbf{r}^{k+1})^T\mathbf{A}\mathbf{p}^k = 0$ (wynika to z poprzedniej iteracji).

Schemat, w którym współczynniki α^k i β^k obliczane są po przez minimalizację residuów nazywamy **uogólnioną metodą najmniejszych residuów** (Generalized Minimal Residual Method — **GMRES**).

Zadania

1. Zmodyfikuj proces iteracji według schematu:
   - oblicz residuum $\mathbf{r}^k = \mathbf{b} - \mathbf{A}\mathbf{x}^k$
   - oblicz poprawkę $\mathbf{p}^k = \mathbf{M}^{-1}\mathbf{r}^k$
   - jeżeli nie jest to pierwsza iteracja: oblicz β^k i nową poprawkę $\mathbf{p}^k = \mathbf{p}^k - \beta^k\mathbf{p}^{k-1}$
   - oblicz α^k
   - wyznacz nowe rozwiązanie $\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha^k\mathbf{p}^k$
   - zachowaj starą poprawkę $\mathbf{p}^k = \mathbf{p}^{k-1}$ (opłaca się też zachować wektor $\mathbf{A}\mathbf{p}^{k-1}$).

Jeśli macierz A jest symetryczna i dodatnio określona. . .

W przypadku naszego zadania MES możemy wykorzystać fakt, że macierz $\mathbf{A}$ jest symetryczna i dodatnio określona. Wtedy zamiast minimalizować $(\mathbf{r}^{k+1})^T\mathbf{r}^{k+1}$ możemy zminimalizować pewien specjalny funkcjonał:

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

Zadania

1. Odwołując się do fizyki naszego przypadku, odpowiedz na poniższe pytania:
   - Czym jest funkcjonał $f(\mathbf{x})$?
   - Dlaczego macierz $\mathbf{A}$ jest symetryczna?
   - Dlaczego $\mathbf{A}$ jest dodatnio określona?

2. Podstaw w powyższym wzorze $\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha^k\mathbf{p}^k$, zróżniczkuj i oblicz α^k. Zauważ, że $\frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{b}^T\mathbf{x} = \text{const.} + \frac{1}{2}(\alpha^k\mathbf{p}^k)^T\mathbf{A}(\alpha^k\mathbf{p}^k) - \mathbf{r}^T(\alpha^k\mathbf{p}^k)$.
3. Analogicznie jak w poprzednim punkcie, podstaw $\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha^k\left(\mathbf{p}^k - \beta^k\mathbf{p}^{k-1}\right)$, zróżniczkuj względem β^k i wyznacz β^k (tym razem $(\mathbf{p}^{k-1})^T\mathbf{r}^k = 0$).

Zadanie

Zastosuj identyczny schemat iteracji jak w poprzednim punkcie ale zmień α^k i β^k. Zbadaj zbieżność.

Taki schemat nazywamy metodą **gradientu sprzężonego** (Conjugate Gradient Method — **CG**).

Uwaga: Aktualnie zbieżność jest bardzo słaba. Wynika to z faktu, że choć $\mathbf{A}$ jest symetryczna to preconditioner z metody Gaussa-Seidla $\mathbf{M}^{-1}$ już nie jest.

Zadanie

Zbadaj zbieżność z preconditionerem diagonalnym, lub wyrażeniem $\mathbf{p}^k = \mathbf{r}^k$ (brakiem preconditionera).

Uwaga: Metodę Conjugate Gradient można zaimplementować w bardziej „zwartej" formie. Taki schemat można znaleźć na Wikipedii, bądź w notatkach z wykładu.