

Wiadomości zebrane –
metody bezgradientowe

Metoda simplexu Neldera i Meada

Metoda simplexu Neldera-Meada jest metodą z grupy bezgradientowych.

Metoda ta polega na utworzeniu w przestrzeni R^{n+1} n -wymiarowego simplexu o $n+1$ wierzchołkach tak, aby można było go wpisać w powierzchnię reprezentującą badaną funkcję celu.

Dla przestrzeni dwuwymiarowej simpleksem jest dowolny trójkąt;

Dla przestrzeni 3 – wymiarowej simplexem jest dowolny czworościan.

Metoda simplexu Neldera i Meada

Jednowymiarowym simplexem jest odcinek o dwóch wierzchołkach, simplexem dwuwymiarowym jest trójkąt i ogólnie simplexem n -wymiarowym o $n+1$ wierzchołkach jest zbiór wszystkich punktów określonych przez wektory:

$$x = \sum_{j=1}^{n+1} x_j e_j \quad \text{gdzie} \quad \sum_{j=1}^{n+1} x_j = 1 \quad \text{oraz} \quad x_j \geq 0$$

Metoda simplexu Neldera i Meada

Czyli mamy do czynienia z wielościanem o $n+1$ wierzchołkach rozpiętym na $n+1$ wektorach bazowych (e_j). Współrzędne punktów simplexu są oznaczone jako x_j

Na początku procedury wylicza się współrzędne punktów wierzchołkowych simplexu P^j

(dla $j = 1 \dots n+1$) przy założeniu pewnej odległości między tymi wierzchołkami (czyli kroku).

W następnych iteracjach dokonuje się przekształceń simplexu aż odległość pomiędzy jego wierzchołkami w pobliżu poszukiwanego ekstremum będzie mniejsza od założonej dokładności obliczeń d .

Jest to jednocześnie kryterium zbieżności tej metody.

Metoda simplexu Neldera i Meada

Metoda zakłada, że wybrany startowy simpleks jest przekształcony za pomocą pewnych elementarnych operacji geometrycznych:

- Odbicia
- Ekspansji
- Zawężenia
- Redukcji

Każda z tych operacji wierzchołek najmniej odpowiedni – w którym f. celu osiąga największą wartość – zastępowany jest przez lepszy czyli wierzchołek dla którego f. Celu osiąga mniejsza wartość. Czyli simpleks zawężany jest do lokalnego minimum.

Metoda simplexu Neldera i Meada

Zanim przejdziemy do właściwej procedury, musimy przeprowadzić analizę i klasyfikację wierzchołków simpleksu, dla których f przyjmuje największe i najmniejsze wartości:

- \min – indeks wierzchołka, dla którego f przyjmuje najmniejszą wartość – $f(p^j)$, $j=0,1,\dots,n$;

$$f(p^{\min}) \leq f(p^j) \text{ dla dowolnego } 0 \leq j \leq n$$

- \max – indeks wierzchołka, dla którego f przyjmuje największą wartość – $f(p^j)$, $j=0,1,\dots,n$;

$$f(p^{\max}) \geq f(p^j) \text{ dla dowolnego } 0 \leq j \leq n$$

- \bar{p} Środek ciężkości wierzchołków simpleksu bez wierzchołka „max”

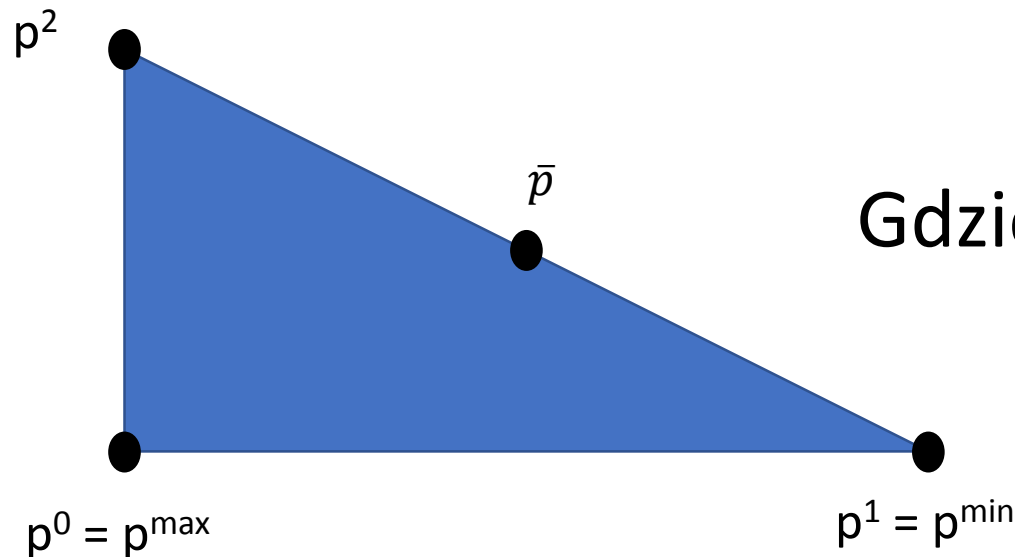
$$\bar{p} = \frac{1}{n} \left(\sum_{j \neq \max} p^j \right)$$

Metoda simplexu Neldera i Meada – przykład 2d

$n=2$, simpleks jest trójkątem ($n+1$ wierzchołków);

Wierzchołki: p^0 , p^1 , p^2 ;

Założmy że $p^1 = p^0 + 2e^1$ oraz $p^2 = p^0 + e^2$ oraz:



Gdzie: $\bar{p} = \frac{1}{2} (p^1 + p^2)$

Metoda simplexu Neldera i Meada – przykład 2d

Po wyznaczeniu p^{\min} , p^{\max} oraz \bar{p} rozpoczynamy procedurę znalezienia minimum f .

W każdej iteracji będą występowały etapy wymienione wcześniej.

- a. Odbicie – pierwszy wyznaczony punkt jest symetrycznym obrazem p^{\max} względem punktu $\bar{p} \rightarrow$ uzyskujemy pkt. Oznaczony p^{odb}

$$p^{\text{odb}} = \bar{p} + \alpha(\bar{p} - p^{\max}) \text{ gdzie } \alpha \in (0,1 >$$

; zwykle $\alpha = 1$

$$p^{\text{odb}} = 2\bar{p} - p^{\max}$$

Metoda simplexu Neldera i Meada – przykład 2d

Po odbiciu możemy spodziewać się następujących przypadków (możliwości):

1. $f(p^{\min}) \leq f(p^{\text{odb}}) < f(p^{\max})$
2. $f(p^{\text{odb}}) < f(p^{\min})$
3. $f(p^{\max}) \leq f(p^{\text{odb}})$

~~Akceptacja odbicia jest warunkowana przez warunek 1.~~

Wtedy zastępujemy p^{\max} przez p^{odb} (nowy simplex tworzą punkty p^{odb} , p^2 , p^{\min}) ; aktualizujemy indeksy min i max oraz sprawdzamy położenie \bar{p}

Metoda simplexu Neldera i Meada – przykład 2d

- b. Ekspansja – następuje po spełnieniu warunku 2 czyli $f(p^{odb}) < f(p^{min})$ co sprowadza się do poszukiwania w kierunku określonym przez odbicie kolejnego punktu (dalej niż p^{odb}), oznacza też brak akceptacji odbicia. Nowy pkt.:

$$p^{eks} = \bar{p} + \gamma(p^{odb} - \bar{p})$$

- Akceptacja ekspansji gdy $f(p^{eks}) < f(p^{odb}) \rightarrow$ zastępujemy p^{max} przez p^{eks} czyli w przykładzie, o którym mowa nowy simpleks tworzą wierzchołki: p^{eks} , p^2 , p^{min} ; aktualizujemy indeksy min i max oraz sprawdzamy położenie \bar{p}
- Brak akceptacji ekspansji gdy $f(p^{eks}) \geq f(p^{odb}) \rightarrow$ zastępujemy p^{max} przez p^{odb} czyli nowy simpleks tworzą wierzchołki: p^{odb} , p^2 , p^{min} ; aktualizujemy indeksy min i max oraz sprawdzamy położenie \bar{p}

Metoda simplexu Neldera i Meada – przykład 2d

c. Zawężenie – dla sytuacji gdy nie akceptujemy odbicia, czyli gdy $f(p^{\text{odb}}) \geq f(p^{\text{max}})$. Wtedy zawężamy simpleks i nowy wierzchołek p^{zaw} wyliczamy według wzoru:

$$p^{\text{zaw}} = \bar{p} + \delta(p^{\text{max}} - \bar{p})$$

Gdzie współczynnik zawężenia $\delta \in (0,1)$

Gdy p^{zaw} daje poprawę czyli $f(p^{\text{zaw}}) < f(p^{\text{max}})$ wtedy p^{max} zastępujemy przez p^{zaw} czyli nowy simpleks tworzą wierzchołki: p^{zaw} , p^2 , p^{min} ; aktualizujemy indeksy min i max i przechodzimy do następnej iteracji

Metoda simplexu Neldera i Meada – przykład 2d

d. Redukcja – gdy spełniona jest nierówność $f(p^{\text{zaw}}) \geq f(p^{\text{max}})$. Wtedy p^{min} pozostaje bez zmian a cały simpleks redukujemy według następującego przepisu:

$$p^j \leftarrow \beta(p^j + p^{\text{min}}), j = 0, \dots, n; \quad j \neq \text{min}$$

Gdzie $\beta \in (0,1)$ i jest określany jako współczynnik redukcji

W następnej iteracji nowy simpleks zostaje utworzony dla punktów określonych według powyższej formuły:
 p^0, \dots, p^n

Metoda simplexu Neldera i Meada – przykład 2d

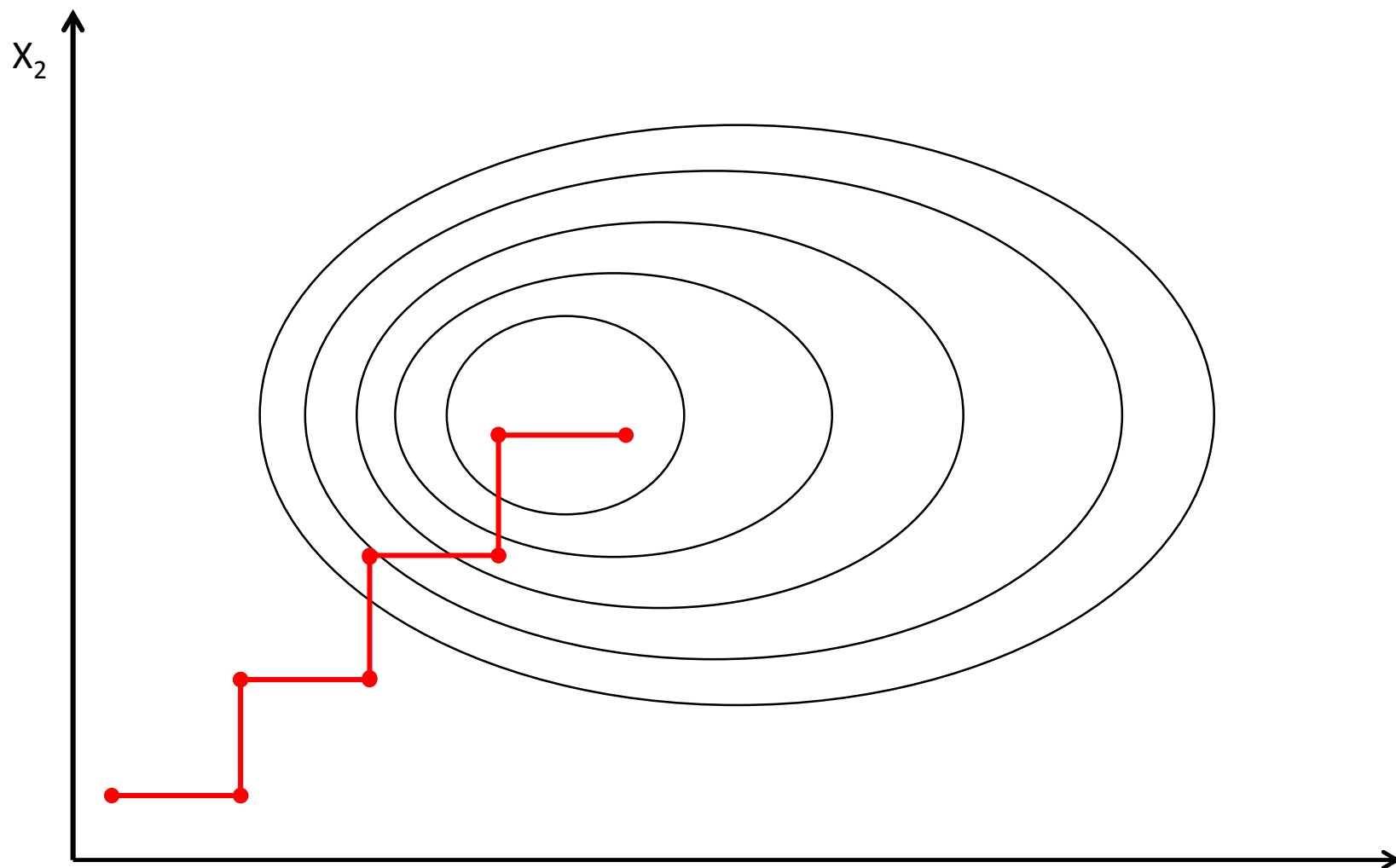
W tej metodzie przyjęte jest geometryczne kryterium stopu – odległości poszczególnych wierzchołków simpleksu są mniejsze od zadanego ε

$$\max_{j=0,\dots,n} \|p^{min} - p^j\| < \varepsilon$$

Metoda spadku względem współrzędnych

- 3) Zmieniamy kolejną współrzędną (y) na tych samych zasadach.
- 4) Obliczenia kończymy, gdy wrócimy po raz drugi do tego samego punktu. Można wówczas przyjąć, iż minimum zostało wyznaczone z dokładnością kroku k .

Metoda spadku względem współrzędnych



Jak widać z rysunku metoda ta jest dość wolno zbieżna, a przy tym stosunkowo mało dokładna (zależy to od długości kroku). W celu zwiększenia dokładności można zmniejszyć krok, lecz to z kolei prowadzi do wydłużenia czasu działania procedury (większa ilość iteracji)

Metoda Gaussa-Seidela

Metoda Gaussa-Seidela, zwana również relaksacyjną jest rozwinięciem metody spadku względem współrzędnych. Również polega ona na minimalizacji funkcji wzdłuż kierunków ortogonalnej bazy $S_1 \dots S_n$. Baza ta utworzona jest z wektorów jednostkowych układu współrzędnych kartezjańskich. Różnica między standardową metodą spadku względem współrzędnych a metodą Gaussa-Seidela polega na tym, iż w **metodzie Gaussa-Seidela krok nie jest stały**.

Metoda Gaussa-Seidela

Ta metoda również więc polega na takim poruszaniu się w dziedzinie funkcji, aby osiągnąć szukane minimum za pomocą zmiany tylko jednej współrzędnej w jednym kroku.

Metoda ta jest szybciej zbieżna od klasycznej metody spadku względem współrzędnych, a przy tym dokładniejsza.

Metoda Gaussa-Seidela

Oznaczenia:

- x_0 - punkt startowy
- $S_1 \dots S_n$ - baza wektorów ortogonalnych (dla dwóch zmiennych $n = 2$)
- e_0 - początkowa długość kroku
- e_j - wymagana dokładność obliczeń w j-tym kierunku
- e - wymagana dokładność obliczeń minimum globalnego

Metoda Gaussa-Seidela

Algorytm:

- 1) W punkcie startowym (x_0, y_0) obliczamy wartość funkcji $f(x_0, y_0)$ - przyjmujemy funkcję dwóch zmiennych dla porównania z metodą spadku względem współrzędnych.
- 2) Obliczamy pochodną funkcji (po długości kroku) w j-tym kierunku i poruszamy się w stronę malejących wartości funkcji. Długość kroku (optymalną) obliczamy z warunku istnienia ekstremum.

$$l \rightarrow \frac{df(l)}{dl} = 0$$

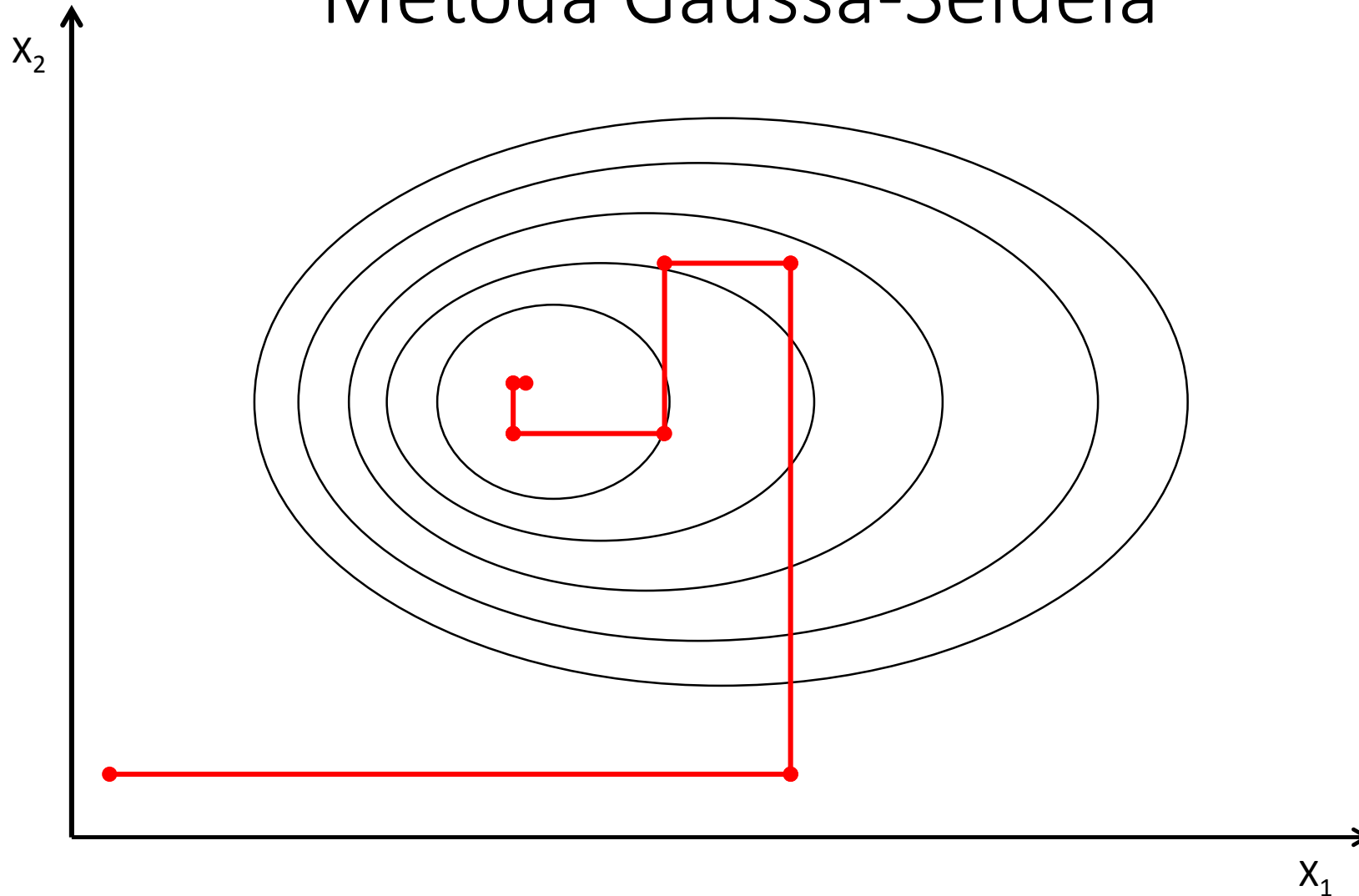
gdzie x_j oznacza kolejne kierunki (w tym przypadku współrzędną x)

Metoda Gaussa-Seidela

3) Zmieniamy kolejną współrzędną (y) na tych samych zasadach (teraz x_j oznacza współrzędną y).

4) Obliczenia kończymy, gdy przesunięcie punktu wzdłuż wszystkich kierunków bazy (w naszym przypadku dwóch) jest mniejsze od wymaganej dokładności.

Metoda Gaussa-Seidela



Algorytm ten jest jednak mało efektywny, gdy poziomice funkcji mają kształt długich wąskich dolin. Wtedy poruszać się musimy małą długością kroku, co zwiększa ilość iteracji, przy czym dla tego rodzaju funkcji często ta metoda okazuje się zawodna.

Metoda Powella

Ta metoda jest stosowana dla funkcji, których poziomice mają kształt wąskich dolin. Można dzięki niej uzyskać znaczną poprawę szybkości zbieżności w stosunku do metody Gaussa-Seidela. Modyfikacja kierunku poszukiwań następuje tu w wyniku wprowadzania do bazy ortogonalnej kierunków sprzężonych do już istniejących. Stosuje się dwa warianty metody Powella.

Metoda Powella

Wariant 1: do istniejącej bazy wprowadza się kierunki sprzężone co obieg (czyli po minimalizacji wzdłuż n kierunków obowiązującej bazy),

Wariant 2: wprowadzamy kierunki sprzężone po spełnieniu określonego warunku.

Ponieważ kierunki wzajemnie sprzężone są liniowo niezależne, w obu wariantach metody Powella zachowany pozostaje warunek jednoznaczności przekształcenia bazy kierunków poszukiwań. Dzięki temu mamy pewność, iż nie nastąpi redukcja wymiarowości bazy, co prowadziłoby do niezbieżności metody.

Metoda Powella

Kierunki e_i oraz e_j są wzajemnie sprzężone względem dodatnio określonej macierzy A , jeżeli zachodzi:

$$e_i A e_j = 0 \text{ dla } i \neq j$$

Metoda Powella

Kryterium zbieżności:

Do ustalenia warunków zakończenia działania procedury iteracyjnej Powell zastosował następujący algorytm:

1) Wykonywanie standardowej procedury aż do momentu, gdy w kolejnej iteracji przesunięcie punktu wzdłuż poszczególnych kierunków poszukiwań będzie mniejsze niż 0,1 wymaganej dokładności obliczeń d . Znaleziony punkt oznaczony jest jako p^a .

Metoda Powella

- 2) Obliczenie nowego punktu startowego przez pomnożenie współrzędnych punktu p^a przez $10d$.
- 3) Powtórzenie czynności z punktu 1). Znaleziony punkt oznaczony jest jako p^b .
- 4) Znalezienie minimum funkcji wzdłuż linii przechodzącej przez oba wyznaczone punkty (p^a i p^b). Znaleziony punkt oznaczony jest jako p^c .
- 5) Zakończenie działania procedury jeżeli $|p^a - p^c|$ oraz $|p^b - p^c|$ są mniejsze od $0,1d$.

Metoda Powella

6) W przeciwnym przypadku wyznaczenie nowego kierunku poszukiwań e_k zgodnie ze wzorem:

$$e_k = \frac{p^a - p^c}{|p^a - p^c|}$$

i włączenie go do bazy na miejsce e_1 , a następnie ponowne przejście do punktu 1)

Metoda Powella

Ze względu na dość dużą złożoność powyższego kryterium w wielu przypadkach **poprzestaje się na sprawdzeniu warunku z punktu 1)** warunek stopu na ilość iteracji.

Istnieje tutaj jednak pewne niebezpieczeństwo popełnienia błędu, gdyż w przypadku silnie nieliniowych funkcji może wystąpić sytuacja, w której przez pewną ilość iteracji przesunięcie punktu wzdłuż kierunków poszukiwań jest znikome, mimo iż obszar ten jest odległy od poszukiwanego minimum.

Metoda Powella - wariant I

Wariant 1 metody Powella został oparty na dwóch twierdzeniach:

1) Jeśli $e_1 \dots e_n$ są wzajemnie sprzężonymi kierunkami względem dodatnio określonej macierzy A i stanowią bazę rozpatrywanej przestrzeni, to startując z dowolnego punktu minimum formy kwadratowej

$$F = a + b^T x + 0,5x^T A x$$

może być wyznaczone w skończonej liczbie iteracji w wyniku minimalizacji funkcji $f(x)$ wzdłuż każdego z tych kierunków e_j tylko raz.

Jeżeli powyższe twierdzenie pozostaje w mocy dla procedury iteracyjnej, to znaczy że posiada ona zbieżność II rzędu.

Metoda Powella - wariant I

2) Jeżeli x_0 jest minimum w kierunku e_i występującym w rozpatrywanej przestrzeni oraz x_1 jest też minimum wzdłuż tego samego kierunku, to kierunek $(x_1 - x_0)$ łączący te dwa minima jest sprzężony z kierunkiem e .

Na początku działania procedury dokonujemy minimalizacji wzdłuż n ortogonalnych kierunków e_1, \dots, e_n (jak w metodzie Gaussa-Seidela), następnie wyznaczamy nowy kierunek sprzężony e_{n+1} zgodnie ze wzorem:

$$e_{n+1} = \frac{x_0 - x}{|x_0 - x|}$$

Metoda Powella - wariant I

i nowy punkt startowy x_0 jako wynik minimalizacji funkcji wzdłuż tego kierunku (e_{n+1}). Potem dokonujemy modyfikacji kierunków usuwając z bazy kierunek e_1 i na jego miejsce włączając kierunek e_{n+1} zmieniając jednocześnie kolejność kierunków (wg reguły $e_k = e_{k+1}$).

Algorytm powtarzamy aż do spełnienia kryterium zbieżności.

Metoda Powella - wariant I

Oznaczenia:

- x_0 - punkt startowy
- $e_1 \dots e_n$ - baza wektorów ortogonalnych (dla dwóch zmiennych $n = 2$)
- l - początkowa długość kroku

Metoda Powella - wariant I

Algorytm:

1) Dla $j = 1 \dots n$ obliczamy l_j minimalizujące $f(x_j)$ oraz współrzędne nowego punktu $x_j = x_{j-1} + l_j e_j$

2) Wyznaczamy składowe kierunku sprzężonego zgodnie ze wzorem:

$$e_{n+1} = \frac{x_0 - x}{|x_0 - x|}$$

3) Określamy l minimalizujące $f(x_{n+1})$ wzdłuż nowego kierunku e_{n+1} oraz wyznaczamy współrzędne nowego punktu startowego

$$x_0 = x_{n+1} = x_n + l e_{n+1}$$

4) Dokonujemy modyfikacji kierunków poszukiwań zgodnie z zasadą $e_r = e_{r+1}$ dla $r = 1 \dots n$
Czynności od kroku 1) do 4) powtarzamy aż spełnione zostanie kryterium na minimum.

Metoda Powella - wariant I

Ze względu na to, iż w pewnych przypadkach procedura nie charakteryzuje się zbieżnością II rzędu, należy przed jej rozpoczęciem dokonać minimalizacji funkcji wzdłuż wszystkich kierunków ortogonalnej bazy początkowej.

Na tym spostrzeżeniu (z dodatkowymi modyfikacjami) została oparta metoda Zangwilla.

Metoda Powella – przykład 2d

