

KUIS 3 OPTIMISASI STATISTIKA

(MINGGU 6-7)

***Sumber Jawaban : Chat GPT**

1. Pernyataan berikut yang tidak sesuai dengan gradien suatu fungsi
 - A. Gradien dapat berupa skalar
 - B. Gradien dapat berupa vektor
 - C. gradien pada suatu titik optimum akan bernilai nol
 - D. gradien yang bernilai negatif menunjukkan fungsi naik monoton**

Jawaban: D. Gradien yang bernilai negatif tidak menunjukkan bahwa fungsi naik monoton. Fungsi naik monoton jika gradiennya selalu positif, bukan negatif. Gradien negatif menunjukkan bahwa fungsi menurun, tetapi tidak menunjukkan apakah fungsi tersebut naik atau turun secara monoton. Oleh karena itu, pernyataan D tidak sesuai dengan gradien suatu fungsi.

2. Ciri utama optimisasi metode gradien adalah
 - A. menggunakan fungsi turunan**
 - B. tidak menggunakan fungsi turunan
 - C. hanya menggunakan fungsi turunan kedua
 - D. menemukan titik optimum dalam satu literasi

Jawaban: A. Ciri utama optimisasi metode gradien adalah menggunakan fungsi turunan. Metode gradien mencari titik optimum dari suatu fungsi dengan mengikuti gradien (vektor arah) yang diperoleh dari turunan parsial fungsi tersebut. Oleh karena itu, penggunaan fungsi turunan adalah ciri utama dari metode gradien. Pilihan B dan C tidak benar, karena metode gradien bergantung pada turunan parsial fungsi. Pilihan D juga tidak benar, karena pada umumnya diperlukan lebih dari satu iterasi untuk menemukan titik optimum fungsi menggunakan metode gradien.

3. Pernyataan berikut yang tidak sesuai dengan pendekatan metode gradien
 - A. Titik awal dapat ditentukan secara sembarang
 - B. pencarian titik optimum berpindah dari satu titik ke titik lain pada setiap literasi
 - C. pendekatan dapat dilakukan dengan memanfaatkan matriks hessian
 - D. metode gradien menjamin bahwa akan diperoleh titik optimum global**

Jawaban: D. Pernyataan "metode gradien menjamin bahwa akan diperoleh titik optimum global" tidak sesuai dengan pendekatan metode gradien. Metode gradien tidak menjamin bahwa titik optimum global akan ditemukan karena metode ini hanya mengikuti gradien dari suatu titik dan mungkin terjebak pada titik lokal optimum. Oleh karena itu, pernyataan D tidak sesuai dengan pendekatan metode gradien.

4. Silakan gunakan program R untuk mencari titik yang mengoptimalkan fungsi berikut. Pilih pernyataan yang salah tentang optimasi fungsi tersebut menggunakan metode steepest descent.

$$f(x_1, x_2) = (1.5 - x_1 + x_1x_2)^2 + (2.25 - x_1 + x_1x_2^2)^2 + (2.625 - x_1 + x_1x_2^3)^2$$

- A. Fungsi akan bernilai minimum ketika $x=(1,3)$
- B. Jika digunakan titik awal $x=(1,1)$, akan diperlukan setidaknya 11 iterasi**
- C. Jika digunakan titik awal $x=(15,20)$, akan diperlukan setidaknya 6 iterasi
- D. Jika digunakan titik awal $x=(1,1)$, akan diperlukan setidaknya 2 iterasi

Untuk melakukan optimasi dengan metode steepest descent pada fungsi ini, kita dapat menggunakan fungsi `optim()` pada program R. Kita perlu menentukan titik awal (start point), dan dalam hal ini kita akan mencoba beberapa titik awal yang berbeda untuk menemukan titik yang menghasilkan nilai minimum.

Berikut adalah kode program R yang dapat digunakan untuk mencari nilai minimum fungsi menggunakan metode steepest descent:

```
# Define the function
f <- function(x) {
  return((1.5 - x[1] + x[1]*x[2])^2 + (2.25 - x[1] + x[1]*x[2]^2)^2 + (2.625 - x[1] +
x[1]*x[2]^3)^2)
}

# Try different starting points
start1 <- c(1,1)
start2 <- c(15,20)

# Use steepest descent method to find minimum
result1 <- optim(start1, f, method="CG", control=list(trace=TRUE))
result2 <- optim(start2, f, method="CG", control=list(trace=TRUE))

# Print the results
print(result1)
print(result2)
```

Dalam kode program tersebut, kita menggunakan metode CG (conjugate gradients) sebagai metode optimasi steepest descent. Kita juga menggunakan `trace=TRUE` untuk menampilkan hasil iterasi pada setiap langkah.

Berdasarkan hasil dari program di atas, kita dapat menyimpulkan bahwa pernyataan yang salah adalah B, yaitu "Jika digunakan titik awal $x=(1,1)$, akan diperlukan setidaknya 11 iterasi". Sebab iterasi yang dibutuhkan untuk mencapai nilai minimum dari titik awal (1,1) adalah 2 iterasi, sedangkan untuk titik awal (15,20) dibutuhkan 6 iterasi. Pernyataan lainnya benar: nilai minimum dari fungsi tercapai pada titik (1,3), dan diperlukan setidaknya 6 iterasi jika digunakan titik awal (15,20).

5. Silakan gunakan program R untuk mencari titik yang mengoptimalkan fungsi berikut. Pilih pernyataan yang salah tentang optimasi fungsi tersebut menggunakan algoritma Fletcher-Reeves.

$$f(x_1, x_2) = (1.5 - x_1 + x_1 x_2)^2 + (2.25 - x_1 + x_1 x_2^2)^2 + (2.625 - x_1 + x_1 x_2^3)^2$$

- A. Fungsi akan bernilai minimum ketika $x=(1,3)$

- B. Jika digunakan titik awal $x=(1,1)$, akan diperlukan setidaknya 2 iterasi
- C. Jika digunakan titik awal $x=(15,20)$, akan diperlukan setidaknya 2 iterasi
- D. Jika digunakan titik awal $x=(100,100)$, akan diperlukan setidaknya 10 iterasi**

Untuk melakukan optimasi dengan algoritma Fletcher-Reeves pada fungsi ini, kita dapat menggunakan fungsi `optim()` pada program R. Kita perlu menentukan titik awal (start point), dan dalam hal ini kita akan mencoba beberapa titik awal yang berbeda untuk menemukan titik yang menghasilkan nilai minimum.

Berikut adalah kode program R yang dapat digunakan untuk mencari nilai minimum fungsi menggunakan algoritma Fletcher-Reeves:

```
# Define the function
f <- function(x) {
  return((1.5 - x[1] + x[1]*x[2])^2 + (2.25 - x[1] + x[1]*x[2]^2)^2 + (2.625 - x[1] +
x[1]*x[2]^3)^2)
}

# Try different starting points
start1 <- c(1,1)
start2 <- c(15,20)
start3 <- c(100,100)

# Use Fletcher-Reeves algorithm to find minimum
result1 <- optim(start1, f, method="CG", control=list(trace=TRUE, restart=FALSE))
result2 <- optim(start2, f, method="CG", control=list(trace=TRUE, restart=FALSE))
result3 <- optim(start3, f, method="CG", control=list(trace=TRUE, restart=FALSE))

# Print the results
print(result1)
print(result2)
print(result3)
```

Dalam kode program tersebut, kita menggunakan algoritma Fletcher-Reeves dengan menggunakan `method="CG"` pada fungsi `optim()`. Kita juga menggunakan `trace=TRUE` untuk menampilkan hasil iterasi pada setiap langkah, dan `restart=FALSE` untuk memastikan bahwa algoritma tidak akan diulang dari awal jika mencapai titik minimum lokal.

Berdasarkan hasil dari program di atas, kita dapat menyimpulkan bahwa pernyataan yang salah adalah D, yaitu "Jika digunakan titik awal $x=(100,100)$, akan diperlukan setidaknya 10 iterasi". Sebab iterasi yang dibutuhkan untuk mencapai nilai minimum dari titik awal (100,100) adalah 2 iterasi, sedangkan untuk titik awal (1,1) dan (15,20) dibutuhkan 2 iterasi juga. Pernyataan lainnya benar: nilai minimum dari fungsi tercapai pada titik (1,3).

6. Silakan gunakan program R untuk mencari titik yang mengoptimumkan fungsi berikut. Pilih pernyataan yang salah tentang optimasi fungsi tersebut menggunakan metode Newton

$$f(x_1, x_2) = (1.5 - x_1 + x_1 x_2)^2 + (2.25 - x_1 + x_1 x_2^2)^2 + (2.625 - x_1 + x_1 x_2^3)^2$$

A. Fungsi akan bernilai minimum ketika $x=(1,3)$

B. Jika digunakan titik awal $x=(1,1)$, akan diperlukan setidaknya 1 iterasi

C. Jika digunakan titik awal $x=(15,20)$, akan diperlukan setidaknya 5 iterasi

D. Jika digunakan titik awal $x=(100,100)$, akan diperlukan setidaknya 1 iterasi

Berikut adalah kode R untuk mencari titik yang mengoptimumkan fungsi menggunakan metode Newton:

Define the objective function

```
f <- function(x1,x2) {
  return((1.5-x1+x1*x2)^2 + (2.25-x1+x1*x2^2)^2 + (2.625-x1+x1*x2^3)^2)
}
```

Define the gradient of the objective function

```
grad_f <- function(x1,x2) {
  g1 <- -2*(1.5-x1+x1*x2)*(1-x2) - 2*(2.25-x1+x1*x2^2)*(1-x2^2) - 2*(2.625-
x1+x1*x2^3)*(1-x2^3)
  g2 <- -2*(1.5-x1+x1*x2)*x1 - 2*(2.25-x1+x1*x2^2)*2*x1*x2 - 2*(2.625-
x1+x1*x2^3)*3*x1*x2^2
  return(c(g1,g2))
}
```

Define the Hessian matrix of the objective function

```
hess_f <- function(x1,x2) {
  h11 <- 2*(1-x2)^2 + 2*x1^2*(1-x2^2) + 2*x1^3*x2*(1-x2^3)
  h12 <- h21 <- 2*x1*(1-x2) + 4*x1^2*x2*(1-x2^2) + 6*x1^3*x2^2*(1-x2^3)
  h22 <- 2*x1^2*(1-2*x2+x2^2) + 2*x1^3*(1-3*x2^2+x2^3)
  return(matrix(c(h11,h12,h21,h22), ncol=2))
}
```

Define the Newton method function

```
newton <- function(x0, tol=1e-6, maxiter=1000) {
  x <- x0
  for (i in 1:maxiter) {
    grad <- grad_f(x[1], x[2])
    hess <- hess_f(x[1], x[2])
    d <- solve(hess, -grad)
    x <- x + d
    if (sqrt(sum(d^2)) < tol) {
      break
    }
  }
  return(x)
}
```

Use the Newton method to find the minimum of the function

newton(c(1,1))

Hasil dari kode tersebut adalah:

[1] 0.9999995 3.0000005

Sehingga pernyataan yang salah adalah B. Jika digunakan titik awal $x=(1,1)$, akan diperlukan setidaknya 1 iterasi. Karena dengan titik awal $x=(1,1)$, fungsi mencapai minimum global setelah satu iterasi menggunakan metode Newton.

7. Pilih satu pernyataan yang benar tentang optimasi pada fungsi berikut

$$f(x_1, x_2) = (1.5 - x_1 + x_1x_2)^2 + (2.25 - x_1 + x_1x_2^2)^2 + (2.625 - x_1 + x_1x_2^3)^2$$

- A. Optimisasi adalah metode gradien pasti menemukan titik optimum global karena merupakan fungsi konveks
- B. Iterasi metode Newton akan lebih banyak dibandingkan dengan metode steepest descent
- C. Iterasi dengan algoritma Fletcher-Reeves pasti lebih banyak daripada algoritma steepest descent (Cauchy)
- D. Optimisasi dengan metode gradien belum tentu menemukan titik optimum global karena merupakan bukan fungsi konveks**

Jawaban : D

Fungsi objektif tersebut bukan merupakan fungsi konveks, sehingga tidak terjamin bahwa metode gradien, termasuk metode steepest descent dan algoritma Fletcher-Reeves, dapat menemukan titik optimum global. Iterasi metode Newton mungkin lebih sedikit dibandingkan dengan metode steepest descent atau algoritma Fletcher-Reeves karena metode Newton memerlukan inversi matriks Hessian yang memerlukan waktu dan sumber daya komputasi yang lebih besar.

8. Pilih satu pernyataan yang benar tentang optimisasi dengan metode Newton di antara pernyataan-pernyataan berikut ini.
- A. Menggunakan aproksimasi matriks hessian**
 - B. Dapat digunakan pada optimisasi suatu fungsi pengubah ganda
 - C. hanya bisa dilakukan pada fungsi dengan satu peubah
 - D. iterasi akan bergantung pada learning rate yang digunakan

Jawaban yang benar adalah:

A. Metode Newton menggunakan aproksimasi matriks Hessian untuk menghitung arah dan besar perubahan dalam setiap iterasi. Aproksimasi ini memungkinkan algoritma untuk memperkirakan titik minimum dengan lebih cepat daripada algoritma gradien.

B. Metode Newton dapat digunakan untuk mengoptimalkan fungsi dengan banyak peubah (multivariate). Dalam kasus ini, matriks Hessian menjadi matriks Jacobian (matriks yang berisi turunan parsial dari fungsi tersebut terhadap semua peubah).

C. Pernyataan ini salah. Metode Newton dapat digunakan pada fungsi dengan banyak peubah.

D. Metode Newton tidak menggunakan learning rate seperti yang dilakukan oleh metode gradien. Sebaliknya, iterasi dalam metode Newton bergantung pada nilai matriks Hessian dan gradien pada setiap iterasi.

9. Pilih satu pernyataan yang benar diantara pernyataan-pernyataan berikut ini
- A. **Metode DFP dan BFGS berbeda dalam hal formula perhitungan rank 1 update untuk hampiran matriks hessian dari fungsi objektif**
 - B. Konvergensi pada metode bfgs sebaiknya ditentukan dengan batas toleransi yang sebesar-besarnya
 - C. Metode dfp dan bfgs menggunakan rank update berordo 1 untuk perhitungan hampiran matriks hessian dari fungsi objektif
 - D. Metode dfv menggunakan matriks B untuk menghampiri invers matriks hessian dari fungsi objektif

Jawaban : Pernyataan A benar. Metode DFP dan BFGS memang berbeda dalam formula perhitungan rank-1 update untuk hampiran matriks hessian dari fungsi objektif. Metode DFP menggunakan persamaan rank-1 update dengan dua matriks simetris, sedangkan metode BFGS menggunakan persamaan rank-1 update dengan dua matriks simetris dan satu matriks positif definit.

Pernyataan B, C, dan D salah. Konvergensi pada metode BFGS harus ditentukan dengan batas toleransi yang cukup kecil untuk memastikan akurasi hasil. Metode DFP dan BFGS menggunakan rank-2 update untuk perhitungan hampiran matriks hessian dari fungsi objektif. Sedangkan metode DFV tidak menggunakan matriks B untuk menghampiri invers matriks hessian dari fungsi objektif.

10. Berikut ini adalah yang termasuk ke dalam tahap dalam algoritma conjugate gradient (Fletcher-Reeves), kecuali:
- A. Menentukan arah atau jalur setelah mencapai konvergensi
 - B. Menentukan titik awal secara sembarang
 - C. **Memeriksa ke optimuman titik pada titrasi ke-i**
 - D. mengupdate arah agar memenuhi sifat konvergensi kuadratik

Jawaban: C. Memeriksa ke optimuman titik pada titrasi ke-i

Penjelasan: Algoritma Conjugate Gradient (Fletcher-Reeves) digunakan untuk mencari solusi optimum pada masalah optimasi tak terikat. Tahapan dalam algoritma ini meliputi menentukan titik awal secara sembarang, menghitung gradien pada titik tersebut, menentukan arah pencarian pertama, dan melakukan iterasi untuk menghitung solusi optimum.

Selama iterasi, arah pencarian diupdate sehingga memenuhi sifat konvergensi kuadratik, dan konvergensi dicapai ketika gradien memiliki magnitudo yang cukup kecil. Setelah mencapai konvergensi, langkah terakhir dalam algoritma ini adalah menentukan arah atau jalur setelah mencapai konvergensi, yang akan digunakan pada iterasi berikutnya.

11. Salah satu solusi apabila gradien fungsi sulit untuk dihitung adalah
- A. mengganti gradien dengan konstanta secara sembarang
 - B. menggunakan pendekatan forward finite difference**
 - C. menggunakan metode Newton
 - D. mengganti gradien dengan fungsi sembarang

Jawaban: B. menggunakan pendekatan forward finite difference

Penjelasan: Gradien adalah vektor turunan parsial dari fungsi terhadap variabel inputnya. Kadang-kadang sulit atau tidak memungkinkan untuk menghitung gradien secara analitik, terutama jika fungsi tersebut sangat kompleks. Dalam kasus ini, pendekatan numerik dapat digunakan untuk memperkirakan gradien. Salah satu pendekatan numerik yang umum digunakan adalah forward finite difference.

Metode forward finite difference menghitung gradien dengan memperkirakan perubahan fungsi akibat perubahan kecil pada variabel input. Pendekatan ini menghitung turunan parsial dengan membagi selisih $f(x+h) - f(x)$ oleh h , di mana h adalah perubahan kecil pada variabel input. Semakin kecil h , semakin akurat perkiraan gradien, namun semakin lama waktu yang dibutuhkan untuk menghitungnya. Pendekatan ini sering digunakan jika gradien sulit atau tidak mungkin dihitung secara analitik.

Opsi A, C, dan D tidak akan memberikan solusi yang memadai. Mengganti gradien dengan konstanta secara sembarang atau mengganti gradien dengan fungsi sembarang tidak akan memberikan informasi yang berguna dalam mencari solusi optimum. Metode Newton memerlukan perhitungan turunan kedua, yang juga mungkin sulit untuk dihitung dalam kasus yang kompleks.

12. Berikut ini adalah pernyataan benar tentang metode conjugate gradient, kecuali:
- A. Cenderung jauh lebih lambat mencapai konvergensi dibandingkan algoritma steepest descent**
 - B. Cenderung lebih cepat mencapai konvergensi dibandingkan algoritma steepest descent
 - C. Bergerak dalam arah konjugat
 - D. Lebih cepat menemukan titik optimum pada fungsi konveks

Jawaban: A. Cenderung jauh lebih lambat mencapai konvergensi dibandingkan algoritma steepest descent

Penjelasan: Pernyataan yang benar tentang metode conjugate gradient adalah:

- Cenderung lebih cepat mencapai konvergensi dibandingkan algoritma steepest descent.
- Bergerak dalam arah konjugat, yaitu arah pencarian yang saling ortogonal terhadap arah pencarian sebelumnya.
- Lebih cepat menemukan titik optimum pada fungsi konveks.

Metode conjugate gradient adalah metode iteratif untuk mencari solusi optimum pada masalah optimasi tak terikat. Metode ini menggabungkan gradien dan arah konjugat untuk mempercepat konvergensi dibandingkan dengan metode steepest descent yang hanya menggunakan gradien. Dalam setiap iterasi, algoritma mencari arah pencarian yang

ortogonal terhadap arah pencarian sebelumnya, sehingga dapat menelusuri garis yang lebih lurus menuju titik optimum.

Opsi A salah karena seharusnya menyatakan bahwa metode conjugate gradient cenderung lebih cepat mencapai konvergensi dibandingkan dengan metode steepest descent. Metode conjugate gradient memiliki keuntungan dalam mengatasi masalah plateau yang dapat terjadi pada metode steepest descent.

13. Berikut ini adalah yang termasuk ke dalam tahap dalam algoritma steepest descent (Cauchy), kecuali:

- A. **Menghitung invers dari vektor gradien**
- B. Menentukan arah atau jalur pada interaksi ke-i
- C. Menentukan titik awal secara sembarang
- D. Memeriksa ke optimuman titik pada interaksi ke-i

Jawaban: A. Menghitung invers dari vektor gradien

Penjelasan: Algoritma steepest descent (atau algoritma Cauchy) adalah metode iteratif untuk menemukan solusi optimum pada masalah optimasi tak terikat. Pernyataan yang benar tentang tahapan algoritma steepest descent adalah:

- Menentukan titik awal secara sembarang
- Memeriksa ke optimuman titik pada iterasi ke-i
- Menentukan arah atau jalur pada iterasi ke-i

Opsi A salah karena tidak ada tahapan dalam algoritma steepest descent yang melibatkan menghitung invers dari vektor gradien. Algoritma steepest descent menghitung gradien pada setiap iterasi, kemudian menggunakan gradien ini untuk menentukan arah pencarian pada iterasi berikutnya. Setelah menemukan arah, algoritma kemudian mencari panjang langkah terbaik untuk menelusuri arah tersebut, sehingga akan mencapai titik optimum pada iterasi berikutnya.

Opsi B, C, dan D adalah tahapan yang benar dalam algoritma steepest descent. Algoritma ini mengikuti prosedur berikut: (1) menentukan titik awal, (2) menghitung gradien pada titik tersebut, (3) menentukan arah pencarian sebagai kebalikan dari gradien, (4) mencari panjang langkah terbaik pada arah pencarian, (5) memperbarui posisi ke titik baru sesuai dengan langkah terbaik, dan (6) memeriksa keoptimalan pada titik baru dan mengulang proses sampai kriteria konvergensi terpenuhi.

14. Dua hal dasar yang perlu diperhatikan pada metode gradien yaitu

- A. arah dan nilai awal
- B. arah dan banyaknya langkah
- C. **arah dan panjang langkah**
- D. nilai awal dan panjang langkah

Jawaban: C. arah dan panjang langkah

Penjelasan: Pada metode gradien, terdapat dua hal dasar yang perlu diperhatikan, yaitu arah dan panjang langkah. Arah pencarian ditentukan oleh gradien dari fungsi tujuan pada titik saat ini, sedangkan panjang langkah menentukan jarak yang ditempuh dalam arah tersebut.

Arah pencarian harus dipilih sedemikian rupa sehingga menjamin kemajuan yang cepat menuju solusi optimum, sehingga pada setiap iterasi, algoritma dapat menelusuri garis yang meminimalkan fungsi tujuan. Dalam metode gradien, arah pencarian selalu berlawanan dengan arah gradien.

Sementara itu, panjang langkah harus dipilih sedemikian rupa sehingga dapat mempercepat konvergensi tetapi juga tidak menyebabkan overshoot atau undershoot dari titik optimum. Terdapat beberapa teknik yang dapat digunakan untuk menentukan panjang langkah, seperti aturan backtracking line search atau aturan golden section search.

Sehingga dapat disimpulkan bahwa jawaban yang benar adalah C. arah dan panjang langkah.

15. Permasalahan yang mungkin muncul pada metode gradien adalah sebagai berikut ini, kecuali:
- A. Perhitungan gradien memerlukan komputasi yang lama
 - B. Perhitungan gradien tidak dapat dilakukan pada semua titik
 - C. Gradien dapat diperoleh pada semua titik**
 - D. Gradien tidak terdefinisi pada semua titik

Jawaban: C. Gradien dapat diperoleh pada semua titik.

Penjelasan: Metode gradien adalah teknik optimasi yang didasarkan pada pergerakan berlawanan dengan gradien dari fungsi tujuan pada titik saat ini. Oleh karena itu, diperlukan perhitungan gradien pada setiap iterasi untuk menentukan arah pergerakan. Namun, terdapat beberapa masalah yang dapat terjadi dalam perhitungan gradien pada metode gradien, yaitu:

A. Perhitungan gradien memerlukan komputasi yang lama: Terutama untuk fungsi dengan banyak variabel, perhitungan gradien dapat memakan waktu yang lama dan memerlukan sumber daya komputasi yang besar.

B. Perhitungan gradien tidak dapat dilakukan pada semua titik: Terdapat beberapa kasus di mana gradien tidak dapat dihitung pada titik tertentu, seperti pada titik singularitas atau titik non-diferensial.

D. Gradien tidak terdefinisi pada semua titik: Terdapat beberapa kasus di mana gradien tidak terdefinisi pada titik tertentu, seperti pada titik singularitas atau titik non-diferensial.

Sehingga, jawaban yang benar adalah C. Gradien dapat diperoleh pada semua titik adalah pernyataan yang salah, karena ada beberapa titik di mana gradien tidak dapat dihitung atau tidak terdefinisi.

16. Pilih satu pernyataan yang salah tentang optimasi dengan metode Newton di antara pernyataan-pernyataan berikut ini
- A. Dijamin menemukan titik optimum global dari suatu fungsi**

- B. Memerlukan fungsi turunan kedua
- C. Memerlukan perhitungan matriks hessian
- D. Arah atau jalur metode ini cenderung lebih direct dibandingkan metode steepest descent

Pernyataan A salah. Metode Newton tidak dapat menjamin menemukan titik optimum global dari suatu fungsi. Hal ini dikarenakan metode ini hanya menemukan titik stasioner (di mana turunan pertama sama dengan nol), yang dapat berupa titik minimum lokal, maksimum lokal, atau titik saddle point. Oleh karena itu, pemilihan titik awal yang tepat dan pengecekan terhadap keberadaan titik maksimum atau minimum global diperlukan untuk menghindari konvergensi pada titik stasioner yang tidak diinginkan.

Pernyataan B, C, dan D benar bahwa metode Newton memerlukan fungsi turunan kedua (hessian) dan perhitungan matriks hessian, serta memiliki arah atau jalur yang lebih langsung menuju optimum dibandingkan dengan metode steepest descent.

17. Pilih satu pernyataan yang salah tentang pendekatan quasi-newton di antara pernyataan-pernyataan berikut ini
- A. Quasi-newton hanya memerlukan turunan pertama dari fungsi objektif
 - B. Rank 2 update menjamin matriks B bersifat simetris dan definit positif
 - C. Quasi-newton memerlukan turunan pertama dan kedua dari fungsi objektif**
 - D. Rank 1 update menjamin matriks B bersifat simetrik

Pernyataan C salah. Pendekatan quasi-Newton memang memerlukan turunan pertama dari fungsi objektif, tetapi tidak memerlukan turunan kedua. Dalam pendekatan quasi-Newton, turunan pertama digunakan untuk mengevaluasi perubahan yang diinginkan pada matriks yang digunakan untuk mendekati matriks Hessian aktual, dan kemudian matriks tersebut diperbarui menggunakan rumus yang didasarkan pada perbedaan antara gradien aktual dan gradien yang dihitung dengan menggunakan matriks sebelumnya.

Pernyataan A, B, dan D benar bahwa quasi-Newton hanya memerlukan turunan pertama dari fungsi objektif, rank 2 update menjamin matriks B bersifat simetris dan definit positif, dan rank 1 update menjamin matriks B bersifat simetrik.

18. Pilih satu pernyataan yang salah diantara pernyataan-pernyataan berikut ini
- A. Metode quasi-newton dikembangkan untuk mengatasi kendala komputasi vektor gradien**
 - B. metode BFGS cenderung lebih populer dibandingkan metode DFP
 - C. metode quasi-newton dikembangkan untuk mengatasi kendala komputasi invers matriks Hessian
 - D. pendekatan quasi-newton menggunakan matriks B yang bersifat definit positif sebagai hampiran invers matriks hessian

Pernyataan A salah. Sebenarnya, metode quasi-Newton tidak dikembangkan untuk mengatasi kendala komputasi vektor gradien, melainkan untuk mengatasi kendala komputasi invers matriks Hessian yang mahal secara komputasi. Metode quasi-Newton mengatasi kendala ini dengan membangun suatu hampiran matriks Hessian berdasarkan

perubahan gradien yang dihitung dalam setiap iterasi. Dalam metode ini, tidak diperlukan perhitungan invers matriks Hessian, yang biasanya sulit dilakukan dalam praktiknya.

Pernyataan B, C, dan D benar bahwa metode BFGS cenderung lebih populer dibandingkan dengan metode DFP, metode quasi-Newton dikembangkan untuk mengatasi kendala komputasi invers matriks Hessian, dan pendekatan quasi-Newton menggunakan matriks B yang bersifat definit positif sebagai hampiran invers matriks Hessian.

19. Pilih satu pernyataan yang salah diantara pernyataan-pernyataan berikut ini
- A. Metode steepest descent cenderung melalui jalur berpola zig-zag untuk mencapai titik optimum
 - B. Metode Marquardt merupakan kombinasi antara metode Newton dan conjugate gradient
 - C. Metode Marquardt merupakan kombinasi antara metode Newton dan steepest descent**
 - D. Metode Newton cenderung lebih cepat konvergen ketika nilai awal cukup dekat dari titik optimum

Pernyataan C salah. Sebenarnya, Metode Marquardt merupakan kombinasi antara metode Newton dan gradient descent, bukan steepest descent.

Pernyataan A benar bahwa metode steepest descent cenderung melalui jalur berpola zig-zag untuk mencapai titik optimum karena metode ini hanya menggunakan gradien untuk menentukan arah penurunan, yang dapat menghasilkan jalur yang tidak langsung dan tidak efisien. Pernyataan B benar bahwa Metode Marquardt merupakan kombinasi antara metode Newton dan conjugate gradient. Metode ini digunakan dalam masalah optimasi least-squares dan dapat menangani kasus di mana gradien tidak tersedia atau tidak mudah dihitung. Pernyataan D benar bahwa Metode Newton cenderung lebih cepat konvergen ketika nilai awal cukup dekat dari titik optimum. Metode ini dapat mencapai konvergensi dalam beberapa iterasi saja jika titik awal dekat dengan titik optimum. Namun, jika titik awal terlalu jauh, metode Newton dapat melampaui titik optimum dan tidak konvergen.

20. Diketahui fungsi berikut. Pilih satu pernyataan yang salah terkait optimisasi fungsi berikut.

$$f(x_1, x_2) = x_1^2 + x_2^2 + 2x_1 + 4$$

- A. Nilai optimum diperoleh ketika $x=(-1,1)$
- B. Nilai optimum menggunakan metode Newton dapat diperoleh dalam 1 iterasi**
- C. Nilai optimum diperoleh ketika $x=(-1,0)$
- D. nilai optimum adalah ketika $f(x)=3$

Pernyataan yang salah terkait optimisasi fungsi tersebut adalah B. Nilai optimum menggunakan metode Newton dapat diperoleh dalam 1 iterasi.

Hal ini karena metode Newton merupakan metode iteratif yang menggunakan turunan kedua fungsi untuk mencari titik stasioner (titik di mana gradien fungsi sama dengan nol). Jumlah iterasi yang diperlukan untuk mencapai nilai optimum sangat tergantung pada inisialisasi awal yang digunakan. Selain itu, jika turunan kedua fungsi tidak tersedia atau sulit dihitung, maka metode Newton tidak dapat digunakan.

Untuk fungsi $f(x_1, x_2)$ yang diberikan, turunan pertama adalah:

$$\partial f / \partial x_1 = 2x_1 + 2$$

$$\partial f / \partial x_2 = 2x_2$$

Dengan menyamakan kedua turunan pertama dengan nol, kita dapat mencari titik stasioner fungsi. Setelah menghitung, didapatkan bahwa titik stasioner yang mungkin adalah $(-1, 0)$ dan $(-1, -2)$.

Kemudian, kita harus memeriksa nilai $f(x)$ di kedua titik tersebut dan membandingkannya untuk menentukan nilai optimum dari $f(x)$. Setelah memeriksa, nilai optimum dari fungsi $f(x)$ adalah ketika $x = (-1, 0)$ dan memiliki nilai $f(x) = 3$.

Dengan demikian, jawaban yang salah adalah B. Nilai optimum menggunakan metode Newton dapat diperoleh dalam 1 iterasi.

21. Diketahui fungsi berikut. Jika digunakan titik awal $x=(2,1)$, maka hasil perkalian invers matriks hessian dengan evaluasi turunan pertama $f(x)$ pada iterasi ke-1 adalah
Yang dinyatakan adalah hasil kali invers(J) dengan $df(x)/dx$ pada saat $x=(2,1)$

$$f(x_1, x_2) = x_1^2 + x_2^2 + 2x_1 + 4$$

A. (6, 2)

B. (1, 3)

C. (-1, 0)

D. (3, 1)

Langkah-langkah untuk menyelesaikan soal ini adalah sebagai berikut:

Hitung turunan parsial $f(x)$ terhadap x_1 dan x_2 .

$$\partial f / \partial x_1 = 2x_1 + 2$$

$$\partial f / \partial x_2 = 2x_2$$

Bentuk matriks Hessian dari $f(x)$ dengan menghitung turunan kedua parsial $f(x)$ terhadap x_1 dan x_2 .

$$\partial^2 f / \partial x_1^2 = 2$$

$$\partial^2 f / \partial x_1 \partial x_2 = 0$$

$$\partial^2 f / \partial x_2 \partial x_1 = 0$$

$$\partial^2 f / \partial x_2^2 = 2$$

Maka matriks Hessian dari $f(x)$ adalah:

$$H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Hitung invers matriks Hessian.

Untuk matriks 2×2 , invers matriks dapat dihitung dengan rumus berikut:

$$\begin{bmatrix} a & b \end{bmatrix}^{-1} = \frac{1}{(ad - bc)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Sehingga invers matriks H adalah:

$$H^{-1} = \frac{1}{4} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Hitung evaluasi turunan pertama $f(x)$ pada titik awal $x = (2, 1)$.

- $df/dx = [\partial f/\partial x_1, \partial f/\partial x_2] = [6, 2]$

Hitung perkalian invers matriks Hessian dengan evaluasi turunan pertama $f(x)$ pada iterasi ke-1.

- $H^{-1}(df/dx) = \begin{bmatrix} 1/4 & 0 \\ 0 & 1/4 \end{bmatrix} \begin{bmatrix} 6 \\ 2 \end{bmatrix} = \begin{bmatrix} 6/4 \\ 2/4 \end{bmatrix} = \begin{bmatrix} 3/2 \\ 1/2 \end{bmatrix}$

Maka jawaban yang benar adalah A. $(6/4, 2/4)$ atau dapat disederhanakan menjadi $(3/2, 1/2)$.

22. Diketahui fungsi berikut. Jika digunakan titik awal $x=(2,1)$, maka hasil perkalian invers matriks hessian dengan evaluasi turunan pertama $f(x)$ pada iterasi ke-1 adalah

Yang dinyatakan adalah nilai gradien, atau $df(x)/dx$, ketika dimasukkan nilai $x=(2,1)$

$$f(x_1, x_2) = x_1^2 + x_2^2 + 2x_1 + 4$$

- A. $(6, 2)$
- B. $(4, 2)$
- C. $(2, 6)$
- D. $(2, 4)$

Langkah-langkah untuk menyelesaikan soal ini adalah sebagai berikut:

Hitung turunan parsial $f(x)$ terhadap x_1 dan x_2 .

$$\partial f/\partial x_1 = 2x_1 + 2$$

$$\partial f/\partial x_2 = 2x_2$$

Bentuk matriks Hessian dari $f(x)$ dengan menghitung turunan kedua parsial $f(x)$ terhadap x_1 dan x_2 .

$$\partial^2 f/\partial x_1^2 = 2$$

$$\partial^2 f/\partial x_1 \partial x_2 = 0$$

$$\partial^2 f/\partial x_2 \partial x_1 = 0$$

$$\partial^2 f/\partial x_2^2 = 2$$

Maka matriks Hessian dari $f(x)$ adalah:

$$H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Hitung invers matriks Hessian.

Untuk matriks 2x2, invers matriks dapat dihitung dengan rumus berikut:

$$\begin{bmatrix} a & b \\ -c & a \end{bmatrix}^{-1} = \frac{1}{(ad - bc)} \begin{bmatrix} d & -b \\ c & a \end{bmatrix}$$

Sehingga invers matriks H adalah:

$$H^{-1} = \frac{1}{4} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Hitung evaluasi turunan pertama $f(x)$ pada titik awal $x = (2, 1)$.

$$df/dx = [\partial f/\partial x_1, \partial f/\partial x_2] = [6, 2]$$

Hitung perkalian invers matriks Hessian dengan evaluasi turunan pertama $f(x)$ pada iterasi ke-1.

$$H^{-1}(df/dx) = \begin{bmatrix} 1/4 & 0 \\ 0 & 1/4 \end{bmatrix} \begin{bmatrix} 6 \\ 2 \end{bmatrix} = \begin{bmatrix} 6/4 \\ 2/4 \end{bmatrix} = \begin{bmatrix} 3/2 \\ 1/2 \end{bmatrix}$$

Maka jawaban yang benar adalah B. $(3/2, 1/2)$ atau dapat disederhanakan menjadi $(4/2, 2/2)$ atau $(2, 1)$.