

## Olinjärt exempel

Det lilla nystartade bryggeriet Frikeller ska designa en ny öl, med en blandning av Amarillo-humle och Citra-humle.

## Olinjärt exempel

Det lilla nystartade bryggeriet Frikeller ska designa en ny öl, med en blandning av Amarillo-humle och Citra-humle.

Låt  $x_1$  stå för andelen Amarillo och  $x_2$  andelen Citra.

## Olinjärt exempel

Det lilla nystartade bryggeriet Frikeller ska designa en ny öl, med en blandning av Amarillo-humle och Citra-humle.

Låt  $x_1$  stå för andelen Amarillo och  $x_2$  andelen Citra.

Efter omfattande undersökningar kommer man fram till att smaken blir bäst om man minimerar följande funktion:  $f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$ .

## Olinjärt exempel

Det lilla nystartade bryggeriet Frikeller ska designa en ny öl, med en blandning av Amarillo-humle och Citra-humle.

Låt  $x_1$  stå för andelen Amarillo och  $x_2$  andelen Citra.

Efter omfattande undersökningar kommer man fram till att smaken blir bäst om man minimerar följande funktion:  $f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$ .

Bivillkor:  $0 \leq x_1 \leq 1$ ,  $0 \leq x_2 \leq 1$ , och  $x_1 + x_2 \leq 1$ .

# Optimalitetsvillkor

En viss punkt är *inte* optimal om vi **vill** och **får** gå till en annan punkt,

## Optimalitetsvillkor

En viss punkt är *inte* optimal om vi **vill** och **får** gå till en annan punkt, dvs. om det finns någon punkt som är **bättre (enligt målfunktionen)** och **tillåten (enligt bivillkoren)**.

## Optimalitetsvillkor

En viss punkt är *inte* optimal om vi **vill** och **får** gå till en annan punkt, dvs. om det finns någon punkt som är **bättre (enligt målfunktionen)** och **tillåten (enligt bivillkoren)**.

En punkt är inte optimal om det finns en **tillåten förbättringsriktning**.

## Optimalitetsvillkor

En viss punkt är *inte* optimal om vi **vill** och **får** gå till en annan punkt, dvs. om det finns någon punkt som är **bättre (enligt målfunktionen)** och **tillåten (enligt bivillkoren)**.

En punkt är inte optimal om det finns en **tillåten förbättringsriktning**.  
(Lokalt kriterium.)



## Optimalitetsvillkor

En viss punkt är *inte* optimal om vi **vill** och **får** gå till en annan punkt, dvs. om det finns någon punkt som är **bättre (enligt målfunktionen)** och **tillåten (enligt bivillkoren)**.

En punkt är inte optimal om det finns en **tillåten förbättringsriktning**. (Lokalt kriterium.)

Gäller omvändningen?

## Optimalitetsvillkor

En viss punkt är *inte* optimal om vi **vill** och **får** gå till en annan punkt, dvs. om det finns någon punkt som är **bättre (enligt målfunktionen)** och **tillåten (enligt bivillkoren)**.

En punkt är inte optimal om det finns en **tillåten förbättringsriktning**. (Lokalt kriterium.)

Gäller omvändningen?

Vi är närsynta! Vad händer utanför vårt synfält?

## Optimalitetsvillkor

En viss punkt är *inte* optimal om vi **vill** och **får** gå till en annan punkt, dvs. om det finns någon punkt som är **bättre (enligt målfunktionen)** och **tillåten (enligt bivillkoren)**.

En punkt är inte optimal om det finns en **tillåten förbättringsriktning**. (Lokalt kriterium.)

Gäller omvändningen?

Vi är närsynta! Vad händer utanför vårt synfält?

Vilka slutsatser kan man dra av lokala egenskaper?

## Optimalitetsvillkor

En viss punkt är *inte* optimal om vi **vill** och **får** gå till en annan punkt, dvs. om det finns någon punkt som är **bättre (enligt målfunktionen)** och **tillåten (enligt bivillkoren)**.

En punkt är inte optimal om det finns en **tillåten förbättringsriktning**.  
(Lokalt kriterium.)

Gäller omvändningen?

Vi är närsynta! Vad händer utanför vårt synfält?

Vilka slutsatser kan man dra av lokala egenskaper?

Är lösningen lokalt optimal?

## Optimalitetsvillkor

En viss punkt är *inte* optimal om vi **vill** och **får** gå till en annan punkt, dvs. om det finns någon punkt som är **bättre (enligt målfunktionen)** och **tillåten (enligt bivillkoren)**.

En punkt är inte optimal om det finns en **tillåten förbättringsriktning**. (Lokalt kriterium.)

Gäller omvändningen?

Vi är närsynta! Vad händer utanför vårt synfält?

Vilka slutsatser kan man dra av lokala egenskaper?

Är lösningen lokalt optimal?

Är lösningen globalt optimal?

## Optimalitetsvillkor

En viss punkt är *inte* optimal om vi **vill** och **får** gå till en annan punkt, dvs. om det finns någon punkt som är **bättre (enligt målfunktionen)** och **tillåten (enligt bivillkoren)**.

En punkt är inte optimal om det finns en **tillåten förbättringsriktning**. (Lokalt kriterium.)

Gäller omvändningen?

Vi är närsynta! Vad händer utanför vårt synfält?

Vilka slutsatser kan man dra av lokala egenskaper?

Är lösningen lokalt optimal?

Är lösningen globalt optimal?

För att kunna säga något säkert behövs exakta matematiska definitioner.

# Optimum?

Se föreläsning 2:

Globalt minimum.

Lokalt minimum.

Konvex mängd.

Snittet av konvexa mängder är konvext.

Konvexkombination.

Extrempunkt.

# Konvex funktion

Konvex funktion:



# Konvex funktion

Konvex funktion: "Glad mun."

## Konvex funktion

Konvex funktion: "Glad mun." Snäll funktion för minimering. (rita)

## Konvex funktion

Konvex funktion: "Glad mun." Snäll funktion för minimering. (rita)  
Linjesegmentet mellan två punkter ligger ovanför funktionen.

## Konvex funktion

Konvex funktion: "Glad mun." Snäll funktion för minimering. (rita)  
Linjesegmentet mellan två punkter ligger ovanför funktionen.

### Definition

En funktion  $f(x)$  är **konvex** om

$$f(\lambda x^{(1)} + (1 - \lambda)x^{(2)}) \leq \lambda f(x^{(1)}) + (1 - \lambda)f(x^{(2)}) \text{ för alla } x^{(1)}, x^{(2)}, 0 \leq \lambda \leq 1.$$

## Konvex funktion

Konvex funktion: "Glad mun." Snäll funktion för minimering. (rita)  
Linjesegmentet mellan två punkter ligger ovanför funktionen.

### Definition

En funktion  $f(x)$  är **konvex** om

$$f(\lambda x^{(1)} + (1 - \lambda)x^{(2)}) \leq \lambda f(x^{(1)}) + (1 - \lambda)f(x^{(2)}) \text{ för alla } x^{(1)}, x^{(2)}, 0 \leq \lambda \leq 1.$$

Exempel:  $f(x) = x_1^2$

## Konvex funktion

Konvex funktion: "Glad mun." Snäll funktion för minimering. (rita)  
Linjesegmentet mellan två punkter ligger ovanför funktionen.

### Definition

En funktion  $f(x)$  är **konvex** om

$$f(\lambda x^{(1)} + (1 - \lambda)x^{(2)}) \leq \lambda f(x^{(1)}) + (1 - \lambda)f(x^{(2)}) \text{ för alla } x^{(1)}, x^{(2)}, 0 \leq \lambda \leq 1.$$

Exempel:  $f(x) = x_1^2$

Exempel:  $f(x) = 3x_1$

## Konvex funktion

Konvex funktion: "Glad mun." Snäll funktion för minimering. (rita)  
Linjesegmentet mellan två punkter ligger ovanför funktionen.

### Definition

En funktion  $f(x)$  är **konvex** om

$$f(\lambda x^{(1)} + (1 - \lambda)x^{(2)}) \leq \lambda f(x^{(1)}) + (1 - \lambda)f(x^{(2)}) \text{ för alla } x^{(1)}, x^{(2)}, 0 \leq \lambda \leq 1.$$

Exempel:  $f(x) = x_1^2$

Exempel:  $f(x) = 3x_1$

Kombinera konvexa funktioner:

## Konvex funktion

Konvex funktion: "Glad mun." Snäll funktion för minimering. (rita)  
Linjesegmentet mellan två punkter ligger ovanför funktionen.

### Definition

En funktion  $f(x)$  är **konvex** om

$$f(\lambda x^{(1)} + (1 - \lambda)x^{(2)}) \leq \lambda f(x^{(1)}) + (1 - \lambda)f(x^{(2)}) \text{ för alla } x^{(1)}, x^{(2)}, 0 \leq \lambda \leq 1.$$

Exempel:  $f(x) = x_1^2$

Exempel:  $f(x) = 3x_1$

Kombinera konvexa funktioner:

- $f_1(x) + f_2(x)$  är konvex om  $f_1(x)$  och  $f_2(x)$  är konvexa.



## Konvex funktion

Konvex funktion: "Glad mun." Snäll funktion för minimering. (rita)  
Linjesegmentet mellan två punkter ligger ovanför funktionen.

### Definition

En funktion  $f(x)$  är **konvex** om

$$f(\lambda x^{(1)} + (1 - \lambda)x^{(2)}) \leq \lambda f(x^{(1)}) + (1 - \lambda)f(x^{(2)}) \text{ för alla } x^{(1)}, x^{(2)}, 0 \leq \lambda \leq 1.$$

Exempel:  $f(x) = x_1^2$

Exempel:  $f(x) = 3x_1$

Kombinera konvexa funktioner:

- $f_1(x) + f_2(x)$  är konvex om  $f_1(x)$  och  $f_2(x)$  är konvexa.
- $kf(x)$  är konvex om  $f(x)$  är konvex och  $k \geq 0$ .

## Konvex funktion

Konvex funktion: "Glad mun." Snäll funktion för minimering. (rita)  
Linjesegmentet mellan två punkter ligger ovanför funktionen.

### Definition

En funktion  $f(x)$  är **konvex** om

$$f(\lambda x^{(1)} + (1 - \lambda)x^{(2)}) \leq \lambda f(x^{(1)}) + (1 - \lambda)f(x^{(2)}) \text{ för alla } x^{(1)}, x^{(2)}, 0 \leq \lambda \leq 1.$$

Exempel:  $f(x) = x_1^2$

Exempel:  $f(x) = 3x_1$

Kombinera konvexa funktioner:

- $f_1(x) + f_2(x)$  är konvex om  $f_1(x)$  och  $f_2(x)$  är konvexa.
- $kf(x)$  är konvex om  $f(x)$  är konvex och  $k \geq 0$ .

Exempel:  $f(x) = 2x_1^2 + 3x_2^2$

## Konvex funktion

Konvex funktion: "Glad mun." Snäll funktion för minimering. (rita)  
Linjesegmentet mellan två punkter ligger ovanför funktionen.

### Definition

En funktion  $f(x)$  är **konvex** om

$$f(\lambda x^{(1)} + (1 - \lambda)x^{(2)}) \leq \lambda f(x^{(1)}) + (1 - \lambda)f(x^{(2)}) \text{ för alla } x^{(1)}, x^{(2)}, 0 \leq \lambda \leq 1.$$

Exempel:  $f(x) = x_1^2$

Exempel:  $f(x) = 3x_1$

Kombinera konvexa funktioner:

- $f_1(x) + f_2(x)$  är konvex om  $f_1(x)$  och  $f_2(x)$  är konvexa.
- $kf(x)$  är konvex om  $f(x)$  är konvex och  $k \geq 0$ .

Exempel:  $f(x) = 2x_1^2 + 3x_2^2$

$$f(x) = 127x_1^2 + 921x_2^2 + 78x_1 + 43x_4 - 12x_5$$

# Differentierbar konvex funktion

Egenvärden hos Hessianen: (jämför med andraderivator i endimensionella fallet)

## Differentierbar konvex funktion

Egenvärden hos Hessianen: (jämför med andraderivator i endimensionella fallet)

Positivt semidefinit matris  $\Leftrightarrow$  egenvärden  $\geq 0$ .

Positivt definit matris  $\Leftrightarrow$  egenvärden  $> 0$ .

## Differentierbar konvex funktion

Egenvärden hos Hessianen: (jämför med andraderivator i endimensionella fallet)

Positivt semidefinit matris  $\Leftrightarrow$  egenvärden  $\geq 0$ .

Positivt definit matris  $\Leftrightarrow$  egenvärden  $> 0$ .

### Sats

Om  $f(x)$  är en två gånger kontinuerligt differentierbar funktion, så är  $f(x)$  konvex om dess hessian  $H(x)$  är positivt semidefinit, och strikt konvex om  $H(x)$  är positivt definit.

## Differentierbar konvex funktion

Egenvärden hos Hessianen: (jämför med andraderivator i endimensionella fallet)

Positivt semidefinit matris  $\Leftrightarrow$  egenvärden  $\geq 0$ .

Positivt definit matris  $\Leftrightarrow$  egenvärden  $> 0$ .

### Sats

Om  $f(x)$  är en två gånger kontinuerligt differentierbar funktion, så är  $f(x)$  konvex om dess hessian  $H(x)$  är positivt semidefinit, och strikt konvex om  $H(x)$  är positivt definit.

Beräkna egenvärden och kolla.

## Differentierbar konvex funktion

Egenvärden hos Hessianen: (jämför med andraderivator i endimensionella fallet)

Positivt semidefinit matris  $\Leftrightarrow$  egenvärden  $\geq 0$ .

Positivt definit matris  $\Leftrightarrow$  egenvärden  $> 0$ .

### Sats

Om  $f(x)$  är en två gånger kontinuerligt differentierbar funktion, så är  $f(x)$  konvex om dess hessian  $H(x)$  är positivt semidefinit, och strikt konvex om  $H(x)$  är positivt definit.

Beräkna egenvärden och kolla.

Metod med underdeterminanter:



## Differentierbar konvex funktion

Egenvärden hos Hessianen: (jämför med andraderivator i endimensionella fallet)

Positivt semidefinit matris  $\Leftrightarrow$  egenvärden  $\geq 0$ .

Positivt definit matris  $\Leftrightarrow$  egenvärden  $> 0$ .

### Sats

Om  $f(x)$  är en två gånger kontinuerligt differentierbar funktion, så är  $f(x)$  konvex om dess hessian  $H(x)$  är positivt semidefinit, och strikt konvex om  $H(x)$  är positivt definit.

Beräkna egenvärden och kolla.

Metod med underdeterminanter:

Om varje ledande underdeterminant är positiv, är alla egenvärden positiva.

## Differentierbar konvex funktion

Egenvärden hos Hessianen: (jämför med andraderivator i endimensionella fallet)

Positivt semidefinit matris  $\Leftrightarrow$  egenvärden  $\geq 0$ .

Positivt definit matris  $\Leftrightarrow$  egenvärden  $> 0$ .

### Sats

Om  $f(x)$  är en två gånger kontinuerligt differentierbar funktion, så är  $f(x)$  konvex om dess hessian  $H(x)$  är positivt semidefinit, och strikt konvex om  $H(x)$  är positivt definit.

Beräkna egenvärden och kolla.

Metod med underdeterminanter:

Om varje ledande underdeterminant är positiv, är alla egenvärden positiva.

En "ledande underdeterminant",  $h_k$ , beräknas med de  $k$  första raderna och kolumnerna,

## Differentierbar konvex funktion

Egenvärden hos Hessianen: (jämför med andraderivator i endimensionella fallet)

Positivt semidefinit matris  $\Leftrightarrow$  egenvärden  $\geq 0$ .

Positivt definit matris  $\Leftrightarrow$  egenvärden  $> 0$ .

### Sats

Om  $f(x)$  är en två gånger kontinuerligt differentierbar funktion, så är  $f(x)$  konvex om dess hessian  $H(x)$  är positivt semidefinit, och strikt konvex om  $H(x)$  är positivt definit.

Beräkna egenvärden och kolla.

Metod med underdeterminanter:

Om varje ledande underdeterminant är positiv, är alla egenvärden positiva.

En "ledande underdeterminant",  $h_k$ , beräknas med de  $k$  första raderna och kolumnerna, dvs. alla kvadratiske delmatriser som börjar i övre vänstra hörnet.

## Differentierbar konvex funktion

Exempel:  $\min x_1^2 + 2x_2^2 - 2x_1x_2$ .

## Differentierbar konvex funktion

Exempel:  $\min x_1^2 + 2x_2^2 - 2x_1x_2$ .

Gradienten är  $\nabla f(x) = \begin{pmatrix} 2x_1 - 2x_2 \\ 4x_2 - 2x_1 \end{pmatrix}$ .

## Differentierbar konvex funktion

Exempel:  $\min x_1^2 + 2x_2^2 - 2x_1x_2$ .

Gradienten är  $\nabla f(x) = \begin{pmatrix} 2x_1 - 2x_2 \\ 4x_2 - 2x_1 \end{pmatrix}$ .

Hessianen är  $H(x) = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$ .

## Differentierbar konvex funktion

Exempel:  $\min x_1^2 + 2x_2^2 - 2x_1x_2$ .

Gradienten är  $\nabla f(x) = \begin{pmatrix} 2x_1 - 2x_2 \\ 4x_2 - 2x_1 \end{pmatrix}$ .

Hessianen är  $H(x) = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$ .

Dess egenvärden är  $3 - \sqrt{5} \approx 0.7639$  och  $3 + \sqrt{5} \approx 5.2361$ , båda positiva,

## Differentierbar konvex funktion

Exempel:  $\min x_1^2 + 2x_2^2 - 2x_1x_2$ .

Gradienten är  $\nabla f(x) = \begin{pmatrix} 2x_1 - 2x_2 \\ 4x_2 - 2x_1 \end{pmatrix}$ .

Hessianen är  $H(x) = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$ .

Dess egenvärden är  $3 - \sqrt{5} \approx 0.7639$  och  $3 + \sqrt{5} \approx 5.2361$ , båda positiva, så Hessianen är positivt definit,



## Differentierbar konvex funktion

Exempel:  $\min x_1^2 + 2x_2^2 - 2x_1x_2$ .

Gradienten är  $\nabla f(x) = \begin{pmatrix} 2x_1 - 2x_2 \\ 4x_2 - 2x_1 \end{pmatrix}$ .

Hessianen är  $H(x) = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$ .

Dess egenvärden är  $3 - \sqrt{5} \approx 0.7639$  och  $3 + \sqrt{5} \approx 5.2361$ , båda positiva, så Hessianen är positivt definit, vilket visar att funktionen  $f(x)$  är konvex.

## Differentierbar konvex funktion

Exempel:  $\min x_1^2 + 2x_2^2 - 2x_1x_2$ .

Gradienten är  $\nabla f(x) = \begin{pmatrix} 2x_1 - 2x_2 \\ 4x_2 - 2x_1 \end{pmatrix}$ .

Hessianen är  $H(x) = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$ .

Dess egenvärden är  $3 - \sqrt{5} \approx 0.7639$  och  $3 + \sqrt{5} \approx 5.2361$ , båda positiva, så Hessianen är positivt definit, vilket visar att funktionen  $f(x)$  är konvex.

Metod med underdeterminanter:

## Differentierbar konvex funktion

Exempel:  $\min x_1^2 + 2x_2^2 - 2x_1x_2$ .

Gradienten är  $\nabla f(x) = \begin{pmatrix} 2x_1 - 2x_2 \\ 4x_2 - 2x_1 \end{pmatrix}$ .

Hessianen är  $H(x) = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$ .

Dess egenvärden är  $3 - \sqrt{5} \approx 0.7639$  och  $3 + \sqrt{5} \approx 5.2361$ , båda positiva, så Hessianen är positivt definit, vilket visar att funktionen  $f(x)$  är konvex.

Metod med underdeterminanter:

$$h_1 = 2 > 0.$$

## Differentierbar konvex funktion

Exempel:  $\min x_1^2 + 2x_2^2 - 2x_1x_2$ .

Gradienten är  $\nabla f(x) = \begin{pmatrix} 2x_1 - 2x_2 \\ 4x_2 - 2x_1 \end{pmatrix}$ .

Hessianen är  $H(x) = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$ .

Dess egenvärden är  $3 - \sqrt{5} \approx 0.7639$  och  $3 + \sqrt{5} \approx 5.2361$ , båda positiva, så Hessianen är positivt definit, vilket visar att funktionen  $f(x)$  är konvex.

Metod med underdeterminanter:

$$h_1 = 2 > 0.$$

$$h_2 = 2 * 4 - (-2) * (-2) = 8 - 4 = 4 > 0.$$

## Differentierbar konvex funktion

Exempel:  $\min x_1^2 + 2x_2^2 - 2x_1x_2$ .

Gradienten är  $\nabla f(x) = \begin{pmatrix} 2x_1 - 2x_2 \\ 4x_2 - 2x_1 \end{pmatrix}$ .

Hessianen är  $H(x) = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$ .

Dess egenvärden är  $3 - \sqrt{5} \approx 0.7639$  och  $3 + \sqrt{5} \approx 5.2361$ , båda positiva, så Hessianen är positivt definit, vilket visar att funktionen  $f(x)$  är konvex.

Metod med underdeterminanter:

$$h_1 = 2 > 0.$$

$$h_2 = 2 * 4 - (-2) * (-2) = 8 - 4 = 4 > 0.$$

Så alla ledande underdeterminanter är positiva,

## Differentierbar konvex funktion

Exempel:  $\min x_1^2 + 2x_2^2 - 2x_1x_2$ .

Gradienten är  $\nabla f(x) = \begin{pmatrix} 2x_1 - 2x_2 \\ 4x_2 - 2x_1 \end{pmatrix}$ .

Hessianen är  $H(x) = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$ .

Dess egenvärden är  $3 - \sqrt{5} \approx 0.7639$  och  $3 + \sqrt{5} \approx 5.2361$ , båda positiva, så Hessianen är positivt definit, vilket visar att funktionen  $f(x)$  är konvex.

Metod med underdeterminanter:

$$h_1 = 2 > 0.$$

$$h_2 = 2 * 4 - (-2) * (-2) = 8 - 4 = 4 > 0.$$

Så alla ledande underdeterminanter är positiva, vilket betyder att alla egenvärden är positiva,

## Differentierbar konvex funktion

Exempel:  $\min x_1^2 + 2x_2^2 - 2x_1x_2$ .

Gradienten är  $\nabla f(x) = \begin{pmatrix} 2x_1 - 2x_2 \\ 4x_2 - 2x_1 \end{pmatrix}$ .

Hessianen är  $H(x) = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$ .

Dess egenvärden är  $3 - \sqrt{5} \approx 0.7639$  och  $3 + \sqrt{5} \approx 5.2361$ , båda positiva, så Hessianen är positivt definit, vilket visar att funktionen  $f(x)$  är konvex.

Metod med underdeterminanter:

$$h_1 = 2 > 0.$$

$$h_2 = 2 * 4 - (-2) * (-2) = 8 - 4 = 4 > 0.$$

Så alla ledande underdeterminanter är positiva, vilket betyder att alla egenvärden är positiva, Hessianen är positivt definit,

## Differentierbar konvex funktion

Exempel:  $\min x_1^2 + 2x_2^2 - 2x_1x_2$ .

Gradienten är  $\nabla f(x) = \begin{pmatrix} 2x_1 - 2x_2 \\ 4x_2 - 2x_1 \end{pmatrix}$ .

Hessianen är  $H(x) = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$ .

Dess egenvärden är  $3 - \sqrt{5} \approx 0.7639$  och  $3 + \sqrt{5} \approx 5.2361$ , båda positiva, så Hessianen är positivt definit, vilket visar att funktionen  $f(x)$  är konvex.

Metod med underdeterminanter:

$$h_1 = 2 > 0.$$

$$h_2 = 2 * 4 - (-2) * (-2) = 8 - 4 = 4 > 0.$$

Så alla ledande underdeterminanter är positiva, vilket betyder att alla egenvärden är positiva, Hessianen är positivt definit, och funktionen  $f(x)$  är konvex.



# Konvex mängd given av bivillkor

Bivillkor definierade av en funktion: (rita)

# Konvex mängd given av bivillkor

Bivillkor definierade av en funktion: (rita)

## Sats

Om  $g(x)$  är en konvex funktion, så ger punkterna som uppfyller  $g(x) \leq b$  en konvex mängd.

# Konvex mängd given av bivillkor

Bivillkor definierade av en funktion: (rita)

## Sats

Om  $g(x)$  är en konvex funktion, så ger punkterna som uppfyller  $g(x) \leq b$  en konvex mängd.

Exempel:  $x_1^2 \leq 3$

# Konvex mängd given av bivillkor

Bivillkor definierade av en funktion: (rita)

## Sats

Om  $g(x)$  är en konvex funktion, så ger punkterna som uppfyller  $g(x) \leq b$  en konvex mängd.

Exempel:  $x_1^2 \leq 3$

$x_1^2 + x_2^2 \leq 17$

# Konvex mängd given av bivillkor

Bivillkor definierade av en funktion: (rita)

## Sats

Om  $g(x)$  är en konvex funktion, så ger punkterna som uppfyller  $g(x) \leq b$  en konvex mängd.

Exempel:  $x_1^2 \leq 3$

$$x_1^2 + x_2^2 \leq 17$$

$$127x_1^2 + 921x_2^2 + 78x_1 + 43x_4 - 12x_5 \leq 473$$

# Konvex mängd given av bivillkor

Bivillkor definierade av en funktion: (rita)

## Sats

Om  $g(x)$  är en konvex funktion, så ger punkterna som uppfyller  $g(x) \leq b$  en konvex mängd.

Exempel:  $x_1^2 \leq 3$

$$x_1^2 + x_2^2 \leq 17$$

$$127x_1^2 + 921x_2^2 + 78x_1 + 43x_4 - 12x_5 \leq 473$$

$$23x_1 + 72x_2 + 55x_3 \leq 17$$

# Konvex mängd given av bivillkor

Bivillkor definierade av en funktion: (rita)

## Sats

Om  $g(x)$  är en konvex funktion, så ger punkterna som uppfyller  $g(x) \leq b$  en konvex mängd.

Exempel:  $x_1^2 \leq 3$

$$x_1^2 + x_2^2 \leq 17$$

$$127x_1^2 + 921x_2^2 + 78x_1 + 43x_4 - 12x_5 \leq 473$$

$$23x_1 + 72x_2 + 55x_3 \leq 17$$

$$23x_1 + 72x_2 + 55x_3 \geq 17$$

# Konvex mängd given av bivillkor

Bivillkor definierade av en funktion: (rita)

## Sats

Om  $g(x)$  är en konvex funktion, så ger punkterna som uppfyller  $g(x) \leq b$  en konvex mängd.

Exempel:  $x_1^2 \leq 3$

$$x_1^2 + x_2^2 \leq 17$$

$$127x_1^2 + 921x_2^2 + 78x_1 + 43x_4 - 12x_5 \leq 473$$

$$23x_1 + 72x_2 + 55x_3 \leq 17$$

$$23x_1 + 72x_2 + 55x_3 \geq 17 \text{ (dvs. } -23x_1 - 72x_2 - 55x_3 \leq -17)$$



# Konvexitetskontroll

# Konvexitetskontroll

- Praktisk konvexitetskontroll av funktion:

# Konvexitetskontroll

- Praktisk konvexitetskontroll av funktion:
  - ▶ Dela upp i konvexa bitar.

# Konvexitetskontroll

- Praktisk konvexitetskontroll av funktion:
  - ▶ Dela upp i konvexa bitar.
  - ▶ Använd satserna.

# Konvexitetskontroll

- Praktisk konvexitetskontroll av funktion:
  - ▶ Dela upp i konvexa bitar.
  - ▶ Använd satserna.
  - ▶ I värsta fall beräkna egenvärden.

# Konvexitetskontroll

- Praktisk konvexitetskontroll av funktion:
  - ▶ Dela upp i konvexa bitar.
  - ▶ Använd satserna.
  - ▶ I värsta fall beräkna egenvärden.
- Kända konvexa funktioner:

# Konvexitetskontroll

- Praktisk konvexitetskontroll av funktion:
  - ▶ Dela upp i konvexa bitar.
  - ▶ Använd satserna.
  - ▶ I värsta fall beräkna egenvärden.
- Kända konvexa funktioner:
  - ▶  $f(x) = a$  (konstant)

# Konvexitetskontroll

- Praktisk konvexitetskontroll av funktion:
  - ▶ Dela upp i konvexa bitar.
  - ▶ Använd satserna.
  - ▶ I värsta fall beräkna egenvärden.
- Kända konvexa funktioner:
  - ▶  $f(x) = a$  (konstant)
  - ▶  $f(x) = |x_j|$  (belopp)



# Konvexitetskontroll

- Praktisk konvexitetskontroll av funktion:
  - ▶ Dela upp i konvexa bitar.
  - ▶ Använd satserna.
  - ▶ I värsta fall beräkna egenvärden.
- Kända konvexa funktioner:
  - ▶  $f(x) = a$  (konstant)
  - ▶  $f(x) = |x_j|$  (belopp)
  - ▶  $f(x) = x_j^2$  (kvadrat)

# Konvexitetskontroll

- Praktisk konvexitetskontroll av funktion:
  - ▶ Dela upp i konvexa bitar.
  - ▶ Använd satserna.
  - ▶ I värsta fall beräkna egenvärden.
- Kända konvexa funktioner:
  - ▶  $f(x) = a$  (konstant)
  - ▶  $f(x) = |x_j|$  (belopp)
  - ▶  $f(x) = x_j^2$  (kvadrat)
  - ▶  $f(x) = \sum_j a_{ij}x_j$  (linjär)

# Konvexitetskontroll

- Praktisk konvexitetskontroll av funktion:
  - ▶ Dela upp i konvexa bitar.
  - ▶ Använd satserna.
  - ▶ I värsta fall beräkna egenvärden.
- Kända konvexa funktioner:
  - ▶  $f(x) = a$  (konstant)
  - ▶  $f(x) = |x_j|$  (belopp)
  - ▶  $f(x) = x_j^2$  (kvadrat)
  - ▶  $f(x) = \sum_j a_{ij}x_j$  (linjär)
- Konvext problem (min): Konvex målfunktion, konvex tillåten mängd.

# Konvexitetskontroll

- Praktisk konvexitetskontroll av funktion:
  - ▶ Dela upp i konvexa bitar.
  - ▶ Använd satserna.
  - ▶ I värsta fall beräkna egenvärden.
- Kända konvexa funktioner:
  - ▶  $f(x) = a$  (konstant)
  - ▶  $f(x) = |x_j|$  (belopp)
  - ▶  $f(x) = x_j^2$  (kvadrat)
  - ▶  $f(x) = \sum_j a_{ij}x_j$  (linjär)
- Konvext problem (min): Konvex målfunktion, konvex tillåten mängd.
- Varför är detta intressant?

# Konvexitetskontroll

- Praktisk konvexitetskontroll av funktion:
  - ▶ Dela upp i konvexa bitar.
  - ▶ Använd satserna.
  - ▶ I värsta fall beräkna egenvärden.
- Kända konvexa funktioner:
  - ▶  $f(x) = a$  (konstant)
  - ▶  $f(x) = |x_j|$  (belopp)
  - ▶  $f(x) = x_j^2$  (kvadrat)
  - ▶  $f(x) = \sum_j a_{ij}x_j$  (linjär)
- Konvext problem (min): Konvex målfunktion, konvex tillåten mängd.
- Varför är detta intressant?

# Konvexitetskontroll

- Praktisk konvexitetskontroll av funktion:
  - ▶ Dela upp i konvexa bitar.
  - ▶ Använd satserna.
  - ▶ I värsta fall beräkna egenvärden.
- Kända konvexa funktioner:
  - ▶  $f(x) = a$  (konstant)
  - ▶  $f(x) = |x_j|$  (belopp)
  - ▶  $f(x) = x_j^2$  (kvadrat)
  - ▶  $f(x) = \sum_j a_{ij}x_j$  (linjär)
- Konvext problem (min): Konvex målfunktion, konvex tillåten mängd.
- Varför är detta intressant?  
Ett lokalt optimum i ett konvext problem är alltid ett globalt optimum!

## Frikeller-exempel

$$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$$

då  $x_1 + x_2 \leq 1$  samt  $0 \leq x_1 \leq 1$  och  $0 \leq x_2 \leq 1$

## Frikeller-exempel

$$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$$

då  $x_1 + x_2 \leq 1$  samt  $0 \leq x_1 \leq 1$  och  $0 \leq x_2 \leq 1$

Målfunktion: Summa av konvexa funktioner  $\Rightarrow$  Konvex.



## Frikeller-exempel

$$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$$

då  $x_1 + x_2 \leq 1$  samt  $0 \leq x_1 \leq 1$  och  $0 \leq x_2 \leq 1$

Målfunktion: Summa av konvexa funktioner  $\Rightarrow$  Konvex.

Bivillkor: Alla linjära  $\Rightarrow$  Konvex tillåtet område.

## Frikeller-exempel

$$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$$

då  $x_1 + x_2 \leq 1$  samt  $0 \leq x_1 \leq 1$  och  $0 \leq x_2 \leq 1$

Målfunktion: Summa av konvexa funktioner  $\Rightarrow$  Konvex.

Bivillkor: Alla linjära  $\Rightarrow$  Konvex tillåtet område.

$\Rightarrow$  Problemet konvext.

# Olinjär optimering utan bivillkor

# Olinjär optimering utan bivillkor

Flervariabelanalyskunskap:

Gradienten,  $\nabla f(x)$ , anger lutningen av funktionen  $f(x)$ ,

# Olinjär optimering utan bivillkor

Flervariabelanalyskunskap:

Gradienten,  $\nabla f(x)$ , anger lutningen av funktionen  $f(x)$ ,  
pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

# Olinjär optimering utan bivillkor

Flervariabelanalyskunskap:

Gradienten,  $\nabla f(x)$ , anger lutningen av funktionen  $f(x)$ ,  
pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Max  $f(x)$ : gå i riktningen  $\nabla f(x)$ .

# Olinjär optimering utan bivillkor

Flervariabelanalyskunskap:

Gradienten,  $\nabla f(x)$ , anger lutningen av funktionen  $f(x)$ ,  
pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Max  $f(x)$ : gå i riktningen  $\nabla f(x)$ . (Min  $f(x)$ ): gå i riktningen

# Olinjär optimering utan bivillkor

Flervariabelanalyskunskap:

Gradienten,  $\nabla f(x)$ , anger lutningen av funktionen  $f(x)$ ,  
pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Max  $f(x)$ : gå i riktningen  $\nabla f(x)$ . (Min  $f(x)$ : gå i riktningen  $-\nabla f(x)$ .)



# Olinjär optimering utan bivillkor

Flervariabelanalyskunskap:

Gradienten,  $\nabla f(x)$ , anger lutningen av funktionen  $f(x)$ ,  
pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Max  $f(x)$ : gå i riktningen  $\nabla f(x)$ . (Min  $f(x)$ : gå i riktningen  $-\nabla f(x)$ .)

Var ligger optimum till min  $f(x)$ ?

# Olinjär optimering utan bivillkor

Flervariabelanalyskunskap:

Gradienten,  $\nabla f(x)$ , anger lutningen av funktionen  $f(x)$ ,  
pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Max  $f(x)$ : gå i riktningen  $\nabla f(x)$ . (Min  $f(x)$ : gå i riktningen  $-\nabla f(x)$ .)

Var ligger optimum till min  $f(x)$ ?

$f(x)$  konvex, differentierbar:

# Olinjär optimering utan bivillkor

Flervariabelanalyskunskap:

Gradienten,  $\nabla f(x)$ , anger lutningen av funktionen  $f(x)$ ,  
pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Max  $f(x)$ : gå i riktningen  $\nabla f(x)$ . (Min  $f(x)$ : gå i riktningen  $-\nabla f(x)$ .)

Var ligger optimum till min  $f(x)$ ?

$f(x)$  konvex, differentierbar:  $\nabla f(\hat{x}) = 0$

# Olinjär optimering utan bivillkor

Flervariabelanalyskunskap:

Gradienten,  $\nabla f(x)$ , anger lutningen av funktionen  $f(x)$ ,  
pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Max  $f(x)$ : gå i riktningen  $\nabla f(x)$ . (Min  $f(x)$ : gå i riktningen  $-\nabla f(x)$ .)

Var ligger optimum till min  $f(x)$ ?

$f(x)$  konvex, differentierbar:  $\nabla f(\hat{x}) = 0 \Rightarrow \hat{x}$  är globalt minimum.

# Olinjär optimering utan bivillkor

Flervariabelanalyskunskap:

Gradienten,  $\nabla f(x)$ , anger lutningen av funktionen  $f(x)$ ,  
pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Max  $f(x)$ : gå i riktningen  $\nabla f(x)$ . (Min  $f(x)$ : gå i riktningen  $-\nabla f(x)$ .)

Var ligger optimum till min  $f(x)$ ?

$f(x)$  konvex, differentierbar:  $\nabla f(\hat{x}) = 0 \Rightarrow \hat{x}$  är globalt minimum.

$f(x)$  ej konvex, men differentierbar:

# Olinjär optimering utan bivillkor

Flervariabelanalyskunskap:

Gradienten,  $\nabla f(x)$ , anger lutningen av funktionen  $f(x)$ ,  
pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Max  $f(x)$ : gå i riktningen  $\nabla f(x)$ . (Min  $f(x)$ : gå i riktningen  $-\nabla f(x)$ .)

Var ligger optimum till min  $f(x)$ ?

$f(x)$  konvex, differentierbar:  $\nabla f(\hat{x}) = 0 \Rightarrow \hat{x}$  är globalt minimum.

$f(x)$  ej konvex, men differentierbar:  $\nabla f(\hat{x}) = 0$

# Olinjär optimering utan bivillkor

Flervariabelanalyskunskap:

Gradienten,  $\nabla f(x)$ , anger lutningen av funktionen  $f(x)$ ,  
pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Max  $f(x)$ : gå i riktningen  $\nabla f(x)$ . (Min  $f(x)$ : gå i riktningen  $-\nabla f(x)$ .)

Var ligger optimum till min  $f(x)$ ?

$f(x)$  konvex, differentierbar:  $\nabla f(\hat{x}) = 0 \Rightarrow \hat{x}$  är globalt minimum.

$f(x)$  ej konvex, men differentierbar:  $\nabla f(\hat{x}) = 0 \Rightarrow$  Lutningen är noll.

# Olinjär optimering utan bivillkor

Flervariabelanalyskunskap:

Gradienten,  $\nabla f(x)$ , anger lutningen av funktionen  $f(x)$ ,  
pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Max  $f(x)$ : gå i riktningen  $\nabla f(x)$ . (Min  $f(x)$ : gå i riktningen  $-\nabla f(x)$ .)

Var ligger optimum till min  $f(x)$ ?

$f(x)$  konvex, differentierbar:  $\nabla f(\hat{x}) = 0 \Rightarrow \hat{x}$  är globalt minimum.

$f(x)$  ej konvex, men differentierbar:  $\nabla f(\hat{x}) = 0 \Rightarrow$  Lutningen är noll.  
 $\hat{x}$  kan vara ett lokalt minimum, ett lokalt maximum eller en sadelpunkt.



# Olinjär optimering utan bivillkor

I en *avtaganderiktning* minskar funktionsvärdet.

I en *ökanderiktning* ökar funktionsvärdet.

# Olinjär optimering utan bivillkor

I en *avtaganderiktning* minskar funktionsvärdet.

I en *ökanderiktning* ökar funktionsvärdet.

En riktning  $d$  är en avtaganderiktning för  $f(x)$  i  $x$  om  $\nabla f(x)^T d < 0$ .

En riktning  $d$  är en ökanderiktning för  $f(x)$  i  $x$  om  $\nabla f(x)^T d > 0$ .

# Olinjär optimering utan bivillkor

I en *avtaganderiktning* minskar funktionsvärdet.

I en *ökanderiktning* ökar funktionsvärdet.

En riktning  $d$  är en avtaganderiktning för  $f(x)$  i  $x$  om  $\nabla f(x)^T d < 0$ .

En riktning  $d$  är en ökanderiktning för  $f(x)$  i  $x$  om  $\nabla f(x)^T d > 0$ .

En ökanderiktning pekar in i samma halvrum som gradienten,  
en avtaganderiktning gör inte det.

# Olinjär optimering utan bivillkor

I en *avtaganderiktning* minskar funktionsvärdet.

I en *ökanderiktning* ökar funktionsvärdet.

En riktning  $d$  är en avtaganderiktning för  $f(x)$  i  $x$  om  $\nabla f(x)^T d < 0$ .

En riktning  $d$  är en ökanderiktning för  $f(x)$  i  $x$  om  $\nabla f(x)^T d > 0$ .

En ökanderiktning pekar in i samma halvrum som gradienten, en avtaganderiktning gör inte det.

Exempel: Är riktningen  $d = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$  en avtaganderiktning till funktionen  $f(x) = x_1^2 + 2x_2^2$  i punkten  $\hat{x} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ ?

# Olinjär optimering utan bivillkor

I en *avtaganderiktning* minskar funktionsvärdet.

I en *ökanderiktning* ökar funktionsvärdet.

En riktning  $d$  är en avtaganderiktning för  $f(x)$  i  $x$  om  $\nabla f(x)^T d < 0$ .

En riktning  $d$  är en ökanderiktning för  $f(x)$  i  $x$  om  $\nabla f(x)^T d > 0$ .

En ökanderiktning pekar in i samma halvrum som gradienten, en avtaganderiktning gör inte det.

Exempel: Är riktningen  $d = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$  en avtaganderiktning till funktionen

$f(x) = x_1^2 + 2x_2^2$  i punkten  $\hat{x} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ ?

$$\nabla f(x) = \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix}.$$

# Olinjär optimering utan bivillkor

I en *avtaganderiktning* minskar funktionsvärdet.

I en *ökanderiktning* ökar funktionsvärdet.

En riktning  $d$  är en avtaganderiktning för  $f(x)$  i  $x$  om  $\nabla f(x)^T d < 0$ .

En riktning  $d$  är en ökanderiktning för  $f(x)$  i  $x$  om  $\nabla f(x)^T d > 0$ .

En ökanderiktning pekar in i samma halvrum som gradienten, en avtaganderiktning gör inte det.

Exempel: Är riktningen  $d = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$  en avtaganderiktning till funktionen

$f(x) = x_1^2 + 2x_2^2$  i punkten  $\hat{x} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ ?

$$\nabla f(x) = \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix}. \quad \nabla f(\hat{x}) = \begin{pmatrix} 4 \\ 12 \end{pmatrix}.$$

## Olinjär optimering utan bivillkor

I en *avtaganderiktning* minskar funktionsvärdet.

I en *ökanderiktning* ökar funktionsvärdet.

En riktning  $d$  är en avtaganderiktning för  $f(x)$  i  $x$  om  $\nabla f(x)^T d < 0$ .

En riktning  $d$  är en ökanderiktning för  $f(x)$  i  $x$  om  $\nabla f(x)^T d > 0$ .

En ökanderiktning pekar in i samma halvrum som gradienten, en avtaganderiktning gör inte det.

Exempel: Är riktningen  $d = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$  en avtaganderiktning till funktionen

$f(x) = x_1^2 + 2x_2^2$  i punkten  $\hat{x} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ ?

$\nabla f(x) = \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix}$ .  $\nabla f(\hat{x}) = \begin{pmatrix} 4 \\ 12 \end{pmatrix}$ . Så  $\nabla f(\hat{x})^T d = 4 - 12 = -8 < 0$ .

## Olinjär optimering utan bivillkor

I en *avtaganderiktning* minskar funktionsvärdet.

I en *ökanderiktning* ökar funktionsvärdet.

En riktning  $d$  är en avtaganderiktning för  $f(x)$  i  $x$  om  $\nabla f(x)^T d < 0$ .

En riktning  $d$  är en ökanderiktning för  $f(x)$  i  $x$  om  $\nabla f(x)^T d > 0$ .

En ökanderiktning pekar in i samma halvrum som gradienten, en avtaganderiktning gör inte det.

Exempel: Är riktningen  $d = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$  en avtaganderiktning till funktionen

$f(x) = x_1^2 + 2x_2^2$  i punkten  $\hat{x} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ ?

$\nabla f(x) = \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix}$ .  $\nabla f(\hat{x}) = \begin{pmatrix} 4 \\ 12 \end{pmatrix}$ . Så  $\nabla f(\hat{x})^T d = 4 - 12 = -8 < 0$ .

Ja, det är en avtaganderiktning.



## Olinjär optimering med bivillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

## Olinjär optimering med bivillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Var ligger optimum?

## Olinjär optimering med bivillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Var ligger optimum?

Inga aktiva bivillkor:

## Olinjär optimering med bivillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Var ligger optimum?

Inga aktiva bivillkor: Som utan bivillkor.  $\nabla f(x) = 0$ .

## Olinjär optimering med bivillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Var ligger optimum?

Inga aktiva bivillkor: Som utan bivillkor.  $\nabla f(x) = 0$ .

Ett eller flera aktiva bivillkor:

## Olinjär optimering med bivillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Var ligger optimum?

Inga aktiva bivillkor: Som utan bivillkor.  $\nabla f(x) = 0$ .

Ett eller flera aktiva bivillkor: Det får inte finnas någon **tillåten förbättringsriktning**.

## Olinjär optimering med bivillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Var ligger optimum?

Inga aktiva bivillkor: Som utan bivillkor.  $\nabla f(x) = 0$ .

Ett eller flera aktiva bivillkor: Det får inte finnas någon **tillåten förbättringsriktning**.

$-\nabla f(x)$  är den mest önskade riktningen (brantaste lutningen).

## Olinjär optimering med bivillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Var ligger optimum?

Inga aktiva bivillkor: Som utan bivillkor.  $\nabla f(x) = 0$ .

Ett eller flera aktiva bivillkor: Det får inte finnas någon **tillåten förbättringsriktning**.

$-\nabla f(x)$  är den mest önskade riktningen (brantaste lutningen).

Alla riktningar  $d$  med  $\nabla f(x)^T d < 0$  är avtaganderiktningar.



## Olinjär optimering med bivillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Var ligger optimum?

Inga aktiva bivillkor: Som utan bivillkor.  $\nabla f(x) = 0$ .

Ett eller flera aktiva bivillkor: Det får inte finnas någon **tillåten förbättringsriktning**.

$-\nabla f(x)$  är den mest önskade riktningen (brantaste lutningen).

Alla riktningar  $d$  med  $\nabla f(x)^T d < 0$  är avtaganderiktningar.

Om bivillkoret  $g_i(x) \leq 0$  är aktivt, får  $g_i(x)$  inte öka.

## Olinjär optimering med bivillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Var ligger optimum?

Inga aktiva bivillkor: Som utan bivillkor.  $\nabla f(x) = 0$ .

Ett eller flera aktiva bivillkor: Det får inte finnas någon **tillåten förbättringsriktning**.

$-\nabla f(x)$  är den mest önskade riktningen (brantaste lutningen).

Alla riktningar  $d$  med  $\nabla f(x)^T d < 0$  är avtaganderiktningar.

Om bivillkoret  $g_i(x) \leq 0$  är aktivt, får  $g_i(x)$  inte öka.

$\nabla g_i(x)$  är då den mest förbjudna riktningen (utåtriktade normalen).

## Olinjär optimering med bivillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Var ligger optimum?

Inga aktiva bivillkor: Som utan bivillkor.  $\nabla f(x) = 0$ .

Ett eller flera aktiva bivillkor: Det får inte finnas någon **tillåten förbättringsriktning**.

$-\nabla f(x)$  är den mest önskade riktningen (brantaste lutningen).

Alla riktningar  $d$  med  $\nabla f(x)^T d < 0$  är avtaganderiktningar.

Om bivillkoret  $g_i(x) \leq 0$  är aktivt, får  $g_i(x)$  inte öka.

$\nabla g_i(x)$  är då den mest förbjudna riktningen (utåtriktade normalen).

Alla riktningar  $d$  med  $\nabla g_i(x)^T d > 0$  är förbjudna.

## Olinjär optimering med bivillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Var ligger optimum?

Inga aktiva bivillkor: Som utan bivillkor.  $\nabla f(x) = 0$ .

Ett eller flera aktiva bivillkor: Det får inte finnas någon **tillåten förbättringsriktning**.

$-\nabla f(x)$  är den mest önskade riktningen (brantaste lutningen).

Alla riktningar  $d$  med  $\nabla f(x)^T d < 0$  är avtaganderiktningar.

Om bivillkoret  $g_i(x) \leq 0$  är aktivt, får  $g_i(x)$  inte öka.

$\nabla g_i(x)$  är då den mest förbjudna riktningen (utåtriktade normalen).

Alla riktningar  $d$  med  $\nabla g_i(x)^T d > 0$  är förbjudna.

Så hur vet man om alla avtaganderiktningar är förbjudna?

## Exempel

Exempel:  $\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

## Exempel

Exempel:  $\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Skriv bivillkoren som  $g_1(x) = x_1 + x_2 - 1 \leq 0$ ,  $g_2(x) = x_2 - 0.4 \leq 0$ .

## Exempel

Exempel:  $\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Skriv bivillkoren som  $g_1(x) = x_1 + x_2 - 1 \leq 0$ ,  $g_2(x) = x_2 - 0.4 \leq 0$ .

Gradienter:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - x_2 - 1 \\ -x_1 + 2x_2 - 1 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

## Exempel

Exempel:  $\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Skriv bivillkoren som  $g_1(x) = x_1 + x_2 - 1 \leq 0$ ,  $g_2(x) = x_2 - 0.4 \leq 0$ .

Gradienter:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - x_2 - 1 \\ -x_1 + 2x_2 - 1 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?



## Exempel

Exempel:  $\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Skriv bivillkoren som  $g_1(x) = x_1 + x_2 - 1 \leq 0$ ,  $g_2(x) = x_2 - 0.4 \leq 0$ .

Gradienter:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - x_2 - 1 \\ -x_1 + 2x_2 - 1 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

Är punkten tillåten?

## Exempel

Exempel:  $\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Skriv bivillkoren som  $g_1(x) = x_1 + x_2 - 1 \leq 0$ ,  $g_2(x) = x_2 - 0.4 \leq 0$ .

Gradienter:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - x_2 - 1 \\ -x_1 + 2x_2 - 1 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

Är punkten tillåten? Kolla:

## Exempel

Exempel:  $\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Skriv bivillkoren som  $g_1(x) = x_1 + x_2 - 1 \leq 0$ ,  $g_2(x) = x_2 - 0.4 \leq 0$ .

Gradienter:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - x_2 - 1 \\ -x_1 + 2x_2 - 1 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

Är punkten tillåten? Kolla:

$$g_1(\hat{x}) = 0.$$

## Exempel

Exempel:  $\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Skriv bivillkoren som  $g_1(x) = x_1 + x_2 - 1 \leq 0$ ,  $g_2(x) = x_2 - 0.4 \leq 0$ .

Gradienter:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - x_2 - 1 \\ -x_1 + 2x_2 - 1 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

Är punkten tillåten? Kolla:

$g_1(\hat{x}) = 0$ . Tillåtet.

## Exempel

Exempel:  $\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Skriv bivillkoren som  $g_1(x) = x_1 + x_2 - 1 \leq 0$ ,  $g_2(x) = x_2 - 0.4 \leq 0$ .

Gradienter:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - x_2 - 1 \\ -x_1 + 2x_2 - 1 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

Är punkten tillåten? Kolla:

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt.

## Exempel

Exempel:  $\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Skriv bivillkoren som  $g_1(x) = x_1 + x_2 - 1 \leq 0$ ,  $g_2(x) = x_2 - 0.4 \leq 0$ .

Gradienter:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - x_2 - 1 \\ -x_1 + 2x_2 - 1 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

Är punkten tillåten? Kolla:

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt.

$g_2(\hat{x}) = 0$ .

## Exempel

Exempel:  $\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Skriv bivillkoren som  $g_1(x) = x_1 + x_2 - 1 \leq 0$ ,  $g_2(x) = x_2 - 0.4 \leq 0$ .

Gradienter:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - x_2 - 1 \\ -x_1 + 2x_2 - 1 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

Är punkten tillåten? Kolla:

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt.

$g_2(\hat{x}) = 0$ . Tillåtet.

## Exempel

Exempel:  $\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Skriv bivillkoren som  $g_1(x) = x_1 + x_2 - 1 \leq 0$ ,  $g_2(x) = x_2 - 0.4 \leq 0$ .

Gradienter:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - x_2 - 1 \\ -x_1 + 2x_2 - 1 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

Är punkten tillåten? Kolla:

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt.

$g_2(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt.



## Exempel

Exempel:  $\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Skriv bivillkoren som  $g_1(x) = x_1 + x_2 - 1 \leq 0$ ,  $g_2(x) = x_2 - 0.4 \leq 0$ .

Gradienter:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - x_2 - 1 \\ -x_1 + 2x_2 - 1 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

Är punkten tillåten? Kolla:

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt.

$g_2(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt.

Punkten är tillåten.

## Exempel

Exempel:  $\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Skriv bivillkoren som  $g_1(x) = x_1 + x_2 - 1 \leq 0$ ,  $g_2(x) = x_2 - 0.4 \leq 0$ .

Gradienter:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - x_2 - 1 \\ -x_1 + 2x_2 - 1 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

Är punkten tillåten? Kolla:

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt.

$g_2(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt.

Punkten är tillåten. Båda bivillkoren aktiva.

## Exempel

Exempel:  $\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Skriv bivillkoren som  $g_1(x) = x_1 + x_2 - 1 \leq 0$ ,  $g_2(x) = x_2 - 0.4 \leq 0$ .

Gradienter:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - x_2 - 1 \\ -x_1 + 2x_2 - 1 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

Är punkten tillåten? Kolla:

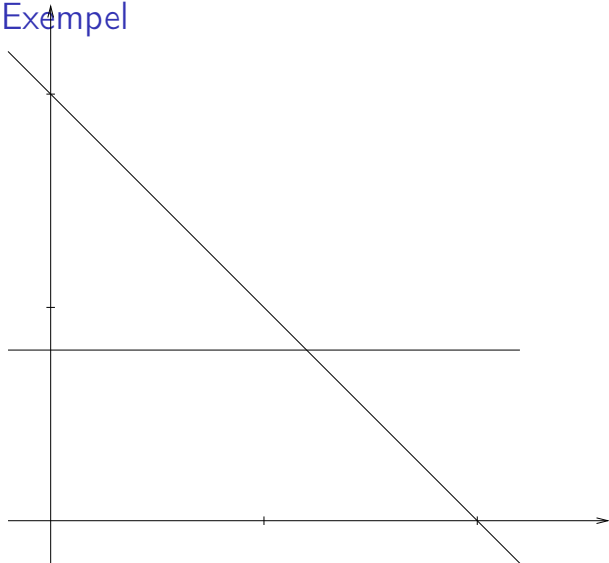
$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt.

$g_2(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt.

Punkten är tillåten. Båda bivillkoren aktiva.

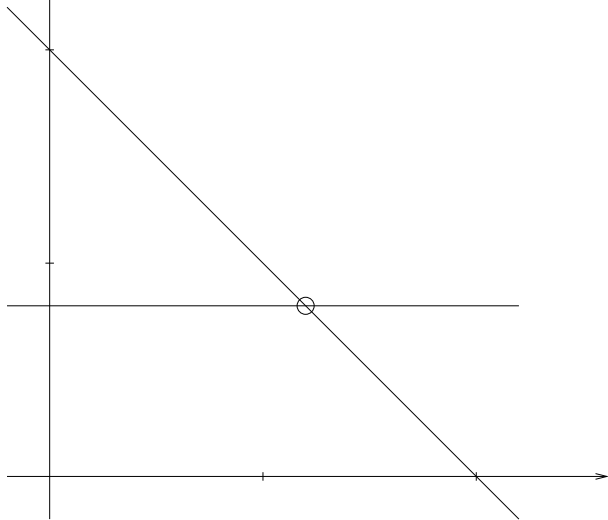
$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}$$

# Exempel



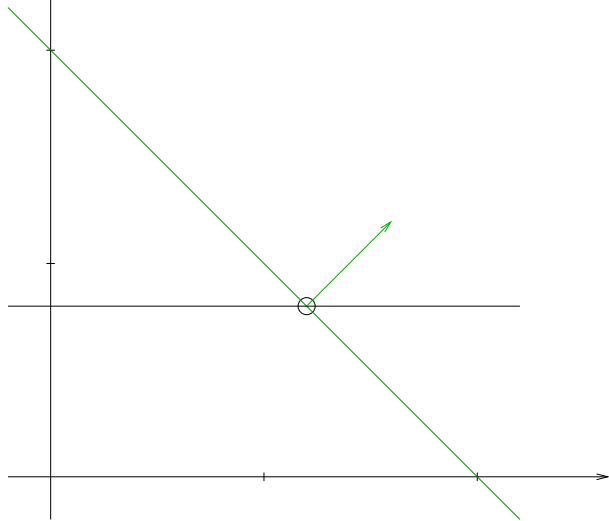
Tillåtet område.

## Exempel



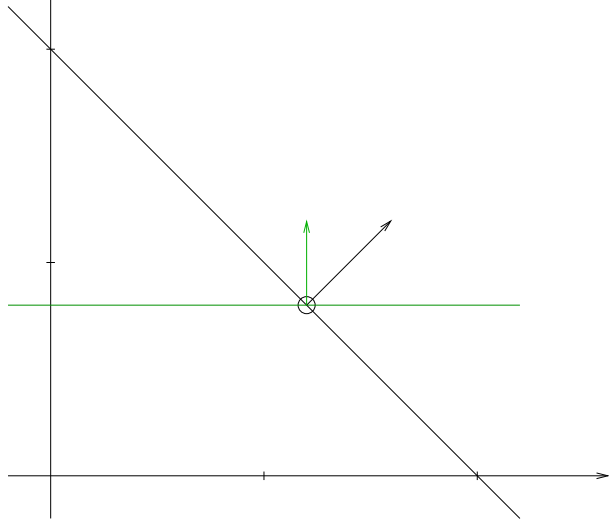
Är denna punkt optimal?

## Exempel



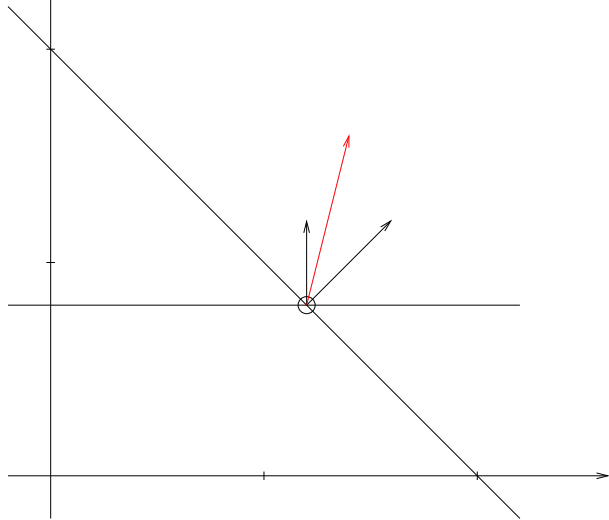
Utåtriktad bivillkorsnormal ( $\nabla g_1(\hat{x})$ ).

## Exempel



Utåtriktad bivillkorsnormal ( $\nabla g_2(\hat{x})$ ).

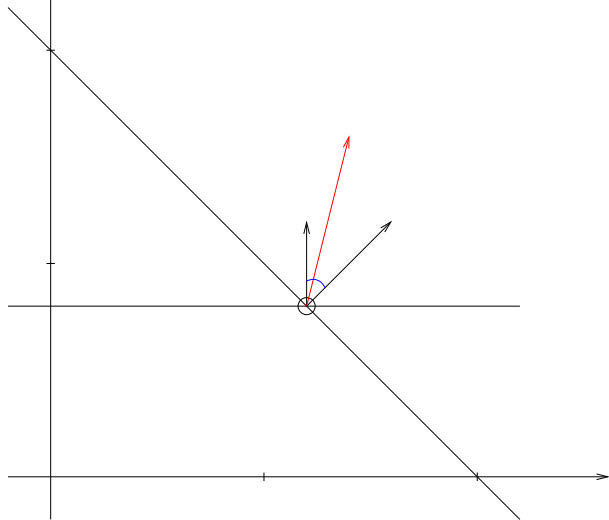
# Exempel



Önskeriktning ( $-\nabla f(\hat{x})$ ).

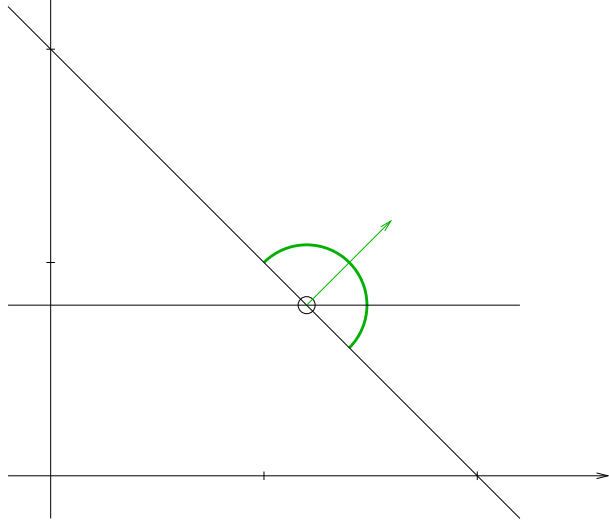


## Exempel



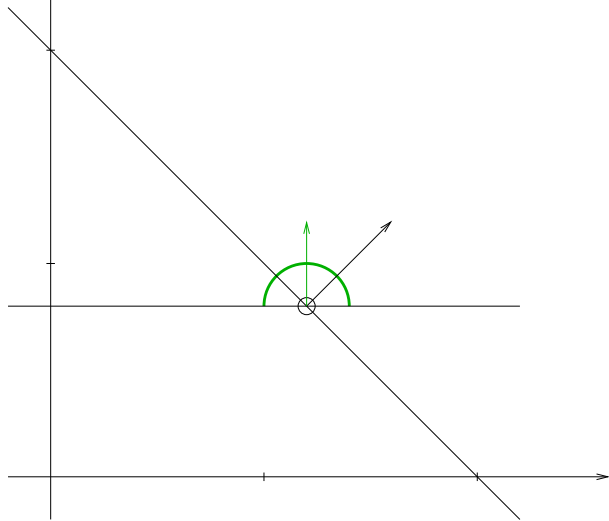
Önskeriktning ( $-\nabla f(\hat{x})$ ). Ligger i konen mellan bivillkorsgradienterna.

## Exempel



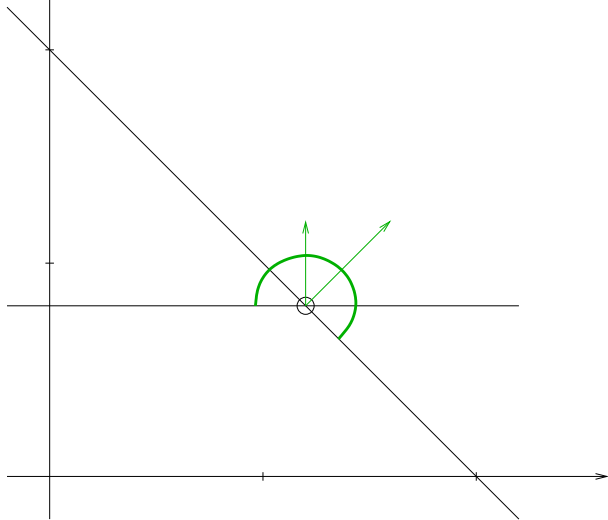
Bivillkor 1 förbjuder dessa riktningar,  $\nabla g_1(\hat{x}) \leq 0$ .

## Exempel



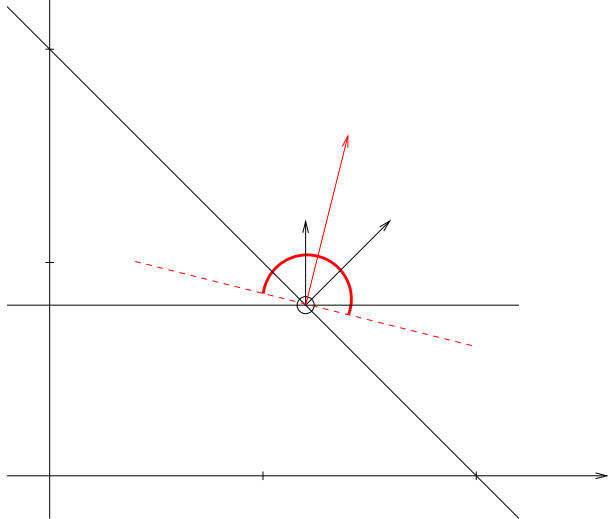
Bivillkor 2 förbjuder dessa riktningar,  $\nabla g_2(\hat{x}) \leq 0$ .

# Exempel



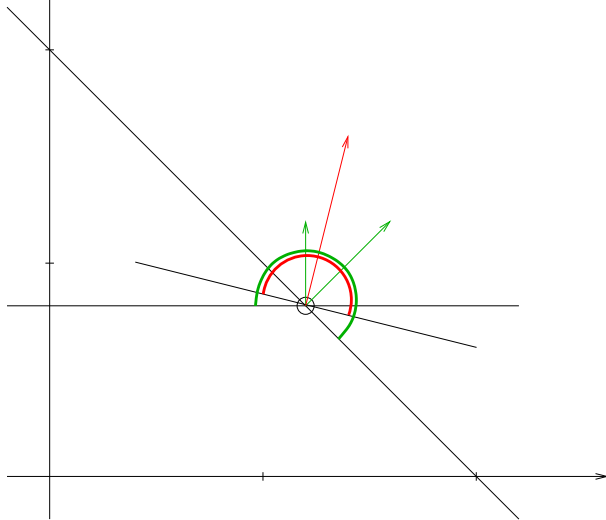
Tillsammans förbjuder bivillkoren dessa riktningar.

## Exempel



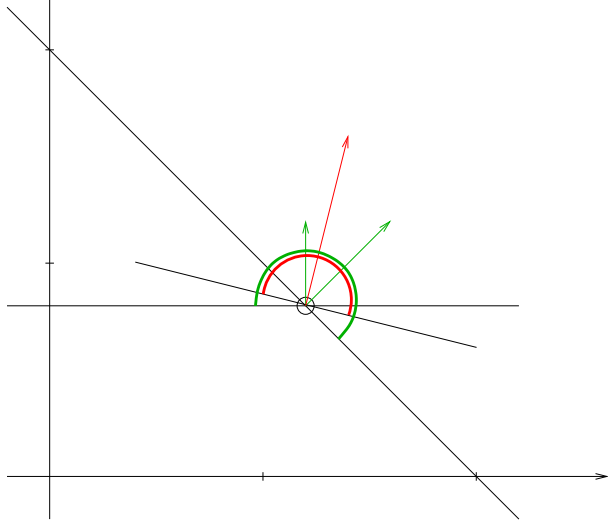
Målfunktionen ger förbättring i dessa riktningar,  $\nabla f(\hat{x}) \leq 0$ .

# Exempel



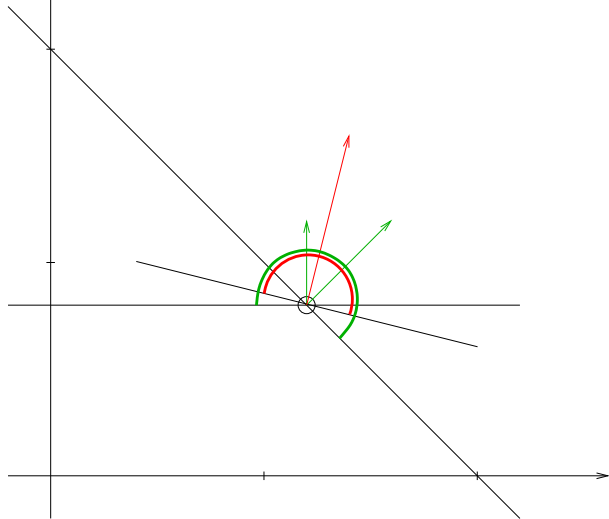
Finns det någon tillåten förbättringsriktning?

## Exempel



Nej, förbättringsriktningarna täcks helt av de förbjudna riktningarna.

# Exempel



Bättre punkt finns inte (i närheten).



# Olinjär optimering med bivillkor: Optimalitetsvillkor

Problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

**Karush-Kuhn-Tuckervillkoren:**

Punkten  $\hat{x}$  är en KKT-punkt om följande är uppfyllt.

**KKT1:**  $g_i(\hat{x}) \leq 0$  för alla  $i$ .

**KKT2:**  $u_i g_i(\hat{x}) = 0$  för alla  $i$ .

**KKT3:**  $\nabla f(\hat{x}) + \sum_{i=1}^m u_i \nabla g_i(\hat{x}) = 0$ .

**KKT4:**  $u_i \geq 0$  för alla  $i$ .

En KKT-punkt är optimal om problemet är konvext.

# Olinjär optimering med bivillkor: Optimalitetsvillkor

Problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

**Karush-Kuhn-Tuckervillkoren:**

Punkten  $\hat{x}$  är en KKT-punkt om följande är uppfyllt.

**KKT1:**  $g_i(\hat{x}) \leq 0$  för alla  $i$ .

**KKT2:**  $u_i g_i(\hat{x}) = 0$  för alla  $i$ .

**KKT3:**  $\nabla f(\hat{x}) + \sum_{i=1}^m u_i \nabla g_i(\hat{x}) = 0$ .

**KKT4:**  $u_i \geq 0$  för alla  $i$ .

En KKT-punkt är optimal om problemet är konvext.

Vad är detta???

# Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Vi ska kontrollera om den givna punkten  $\hat{x}$  är en KKT-punkt.

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Vi ska kontrollera om den givna punkten  $\hat{x}$  är en KKT-punkt.

**KKT1:** Kolla att  $g_i(\hat{x}) \leq 0$  för alla  $i$ .

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Vi ska kontrollera om den givna punkten  $\hat{x}$  är en KKT-punkt.

**KKT1:** Kolla att  $g_i(\hat{x}) \leq 0$  för alla  $i$ .

Kontrollera att punkten är tillåten.

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Vi ska kontrollera om den givna punkten  $\hat{x}$  är en KKT-punkt.

**KKT1:** Kolla att  $g_i(\hat{x}) \leq 0$  för alla  $i$ .

Kontrollera att punkten är tillåten.

**KKT2:** Se till att  $u_i g_i(\hat{x}) = 0$  för alla  $i$  genom att sätta vissa  $u_i = 0$ .

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Vi ska kontrollera om den givna punkten  $\hat{x}$  är en KKT-punkt.

**KKT1:** Kolla att  $g_i(\hat{x}) \leq 0$  för alla  $i$ .

Kontrollera att punkten är tillåten.

**KKT2:** Se till att  $u_i g_i(\hat{x}) = 0$  för alla  $i$  genom att sätta vissa  $u_i = 0$ .

Ska göra en projektion av målfunktionen på de **aktiva** bivillkoren.

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Vi ska kontrollera om den givna punkten  $\hat{x}$  är en KKT-punkt.

**KKT1:** Kolla att  $g_i(\hat{x}) \leq 0$  för alla  $i$ .

Kontrollera att punkten är tillåten.

**KKT2:** Se till att  $u_i g_i(\hat{x}) = 0$  för alla  $i$  genom att sätta vissa  $u_i = 0$ .

Ska göra en projektion av målfunktionen på de **aktiva** bivillkoren.

För att bara få med de aktiva bivillkoren, sätts  $u_i = 0$  för alla icke aktiva bivillkor,



## Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

Vi ska kontrollera om den givna punkten  $\hat{x}$  är en KKT-punkt.

**KKT1:** Kolla att  $g_i(\hat{x}) \leq 0$  för alla  $i$ .

Kontrollera att punkten är tillåten.

**KKT2:** Se till att  $u_i g_i(\hat{x}) = 0$  för alla  $i$  genom att sätta vissa  $u_i = 0$ .

Ska göra en projektion av målfunktionen på de **aktiva** bivillkoren.

För att bara få med de aktiva bivillkoren, sätts  $u_i = 0$  för alla icke aktiva bivillkor, dvs. de som har  $g_i(\hat{x}) < 0$ .

# Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{aligned} \min \quad & f(x) \\ \text{då} \quad & g_i(x) \leq 0 \quad i = 1, \dots, m \end{aligned}$$

Vi ska kontrollera om den givna punkten  $\hat{x}$  är en KKT-punkt.

**KKT1:** Kolla att  $g_i(\hat{x}) \leq 0$  för alla  $i$ .

Kontrollera att punkten är tillåten.

**KKT2:** Se till att  $u_i g_i(\hat{x}) = 0$  för alla  $i$  genom att sätta vissa  $u_i = 0$ .

Ska göra en projektion av målfunktionen på de **aktiva** bivillkoren.

För att bara få med de aktiva bivillkoren, sätts  $u_i = 0$  för alla icke aktiva bivillkor, dvs. de som har  $g_i(\hat{x}) < 0$ .

(I två dimensioner kan man se detta grafiskt.)

# Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

KKT1 och KKT2 ser till att punkten är tillåten, och att  $u_i = 0$  för alla icke aktiva bivillkor.

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

KKT1 och KKT2 ser till att punkten är tillåten, och att  $u_i = 0$  för alla icke aktiva bivillkor.

Gör nu projektionen, dvs. beräkna resterande  $u_i$ :

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

KKT1 och KKT2 ser till att punkten är tillåten, och att  $u_i = 0$  för alla icke aktiva bivillkor.

Gör nu projektionen, dvs. beräkna resterande  $u_i$ :

$$\text{KKT3: } \nabla f(\hat{x}) + \sum_{i=1}^m u_i \nabla g_i(\hat{x}) = 0.$$

# Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

KKT1 och KKT2 ser till att punkten är tillåten, och att  $u_i = 0$  för alla icke aktiva bivillkor.

Gör nu projektionen, dvs. beräkna resterande  $u_j$ :

$$\text{KKT3: } \nabla f(\hat{x}) + \sum_{i=1}^m u_i \nabla g_i(\hat{x}) = 0.$$

$$\text{som är samma som } -\nabla f(\hat{x}) = \sum_i u_i \nabla g_i(\hat{x}).$$

# Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

KKT1 och KKT2 ser till att punkten är tillåten, och att  $u_i = 0$  för alla icke aktiva bivillkor.

Gör nu projektionen, dvs. beräkna resterande  $u_i$ :

$$\text{KKT3: } \nabla f(\hat{x}) + \sum_{i=1}^m u_i \nabla g_i(\hat{x}) = 0.$$

$$\text{som är samma som } -\nabla f(\hat{x}) = \sum_i u_i \nabla g_i(\hat{x}).$$

Observera att  $-\nabla f(\hat{x})$  är vår önskeriktning,



# Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{aligned} \min \quad & f(x) \\ \text{då} \quad & g_i(x) \leq 0 \quad i = 1, \dots, m \end{aligned}$$

KKT1 och KKT2 ser till att punkten är tillåten, och att  $u_i = 0$  för alla icke aktiva bivillkor.

Gör nu projektionen, dvs. beräkna resterande  $u_i$ :

$$\text{KKT3: } \nabla f(\hat{x}) + \sum_{i=1}^m u_i \nabla g_i(\hat{x}) = 0.$$

$$\text{som är samma som } -\nabla f(\hat{x}) = \sum_i u_i \nabla g_i(\hat{x}).$$

Observera att  $-\nabla f(\hat{x})$  är vår önskeriktning, medan  $\nabla g_i(\hat{x})$  är utåtriktade normaler till bivillkoren,

# Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

KKT1 och KKT2 ser till att punkten är tillåten, och att  $u_i = 0$  för alla icke aktiva bivillkor.

Gör nu projektionen, dvs. beräkna resterande  $u_i$ :

$$\text{KKT3: } \nabla f(\hat{x}) + \sum_{i=1}^m u_i \nabla g_i(\hat{x}) = 0.$$

$$\text{som är samma som } -\nabla f(\hat{x}) = \sum_i u_i \nabla g_i(\hat{x}).$$

Observera att  $-\nabla f(\hat{x})$  är vår önskeriktning, medan  $\nabla g_i(\hat{x})$  är utåtriktade normaler till bivillkoren, dvs. de mest förbjudna riktningarna.

# Olinjär optimering med bivillkor: Optimalitetsvillkor

Analys av KKT-villkoren för problemet:

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

KKT1 och KKT2 ser till att punkten är tillåten, och att  $u_i = 0$  för alla icke aktiva bivillkor.

Gör nu projektionen, dvs. beräkna resterande  $u_i$ :

$$\text{KKT3: } \nabla f(\hat{x}) + \sum_{i=1}^m u_i \nabla g_i(\hat{x}) = 0.$$

$$\text{som är samma som } -\nabla f(\hat{x}) = \sum_i u_i \nabla g_i(\hat{x}).$$

Observera att  $-\nabla f(\hat{x})$  är vår önskeriktning, medan  $\nabla g_i(\hat{x})$  är utåtriktade normaler till bivillkoren, dvs. de mest förbjudna riktningarna.

Vi skriver önskeriktningen som en kombination av förbjudna riktningar.

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Vi har räknat ut  $u$  så att 
$$-\nabla f(\hat{x}) = \sum_i u_i \nabla g_i(\hat{x}).$$

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Vi har räknat ut  $u$  så att 
$$-\nabla f(\hat{x}) = \sum_i u_i \nabla g_i(\hat{x}).$$

Påstående: Om  $u_i \geq 0$  för alla  $i$ , så är alla förbättringsriktningar otillåtna.

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Vi har räknat ut  $u$  så att 
$$-\nabla f(\hat{x}) = \sum_i u_i \nabla g_i(\hat{x}).$$

Påstående: Om  $u_i \geq 0$  för alla  $i$ , så är alla förbättringsriktningar otillåtna.

Multiplitera med en valfri riktning  $d$ : 
$$-\nabla f(\hat{x})^T d = \sum_i u_i \nabla g_i(\hat{x})^T d.$$

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Vi har räknat ut  $u$  så att 
$$-\nabla f(\hat{x}) = \sum_i u_i \nabla g_i(\hat{x}).$$

Påstående: Om  $u_i \geq 0$  för alla  $i$ , så är alla förbättringsriktningar otillåtna.

Multiplitera med en valfri riktning  $d$ : 
$$-\nabla f(\hat{x})^T d = \sum_i u_i \nabla g_i(\hat{x})^T d.$$

Om  $d$  är en **tillåten riktning**, så är  $\nabla g_i(\hat{x})^T d \leq 0$  för alla  $i$ .

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Vi har räknat ut  $u$  så att 
$$-\nabla f(\hat{x}) = \sum_i u_i \nabla g_i(\hat{x}).$$

Påstående: Om  $u_i \geq 0$  för alla  $i$ , så är alla förbättringsriktningar otillåtna.

Multiplitera med en valfri riktning  $d$ : 
$$-\nabla f(\hat{x})^T d = \sum_i u_i \nabla g_i(\hat{x})^T d.$$

Om  $d$  är en **tillåten riktning**, så är  $\nabla g_i(\hat{x})^T d \leq 0$  för alla  $i$ .

Om  $u_i \geq 0$  för alla  $i$ , så blir 
$$\sum_i u_i \nabla g_i(\hat{x})^T d \leq 0,$$



## Olinjär optimering med bivillkor: Optimalitetsvillkor

Vi har räknat ut  $u$  så att 
$$-\nabla f(\hat{x}) = \sum_i u_i \nabla g_i(\hat{x}).$$

Påstående: Om  $u_i \geq 0$  för alla  $i$ , så är alla förbättringsriktningar otillåtna.

Multipluera med en valfri riktning  $d$ : 
$$-\nabla f(\hat{x})^T d = \sum_i u_i \nabla g_i(\hat{x})^T d.$$

Om  $d$  är en **tillåten riktning**, så är  $\nabla g_i(\hat{x})^T d \leq 0$  för alla  $i$ .

Om  $u_i \geq 0$  för alla  $i$ , så blir 
$$\sum_i u_i \nabla g_i(\hat{x})^T d \leq 0,$$

vilket betyder att  $-\nabla f(\hat{x})^T d \leq 0$ ,

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Vi har räknat ut  $u$  så att 
$$-\nabla f(\hat{x}) = \sum_i u_i \nabla g_i(\hat{x}).$$

Påstående: Om  $u_i \geq 0$  för alla  $i$ , så är alla förbättringsriktningar otillåtna.

Multiplitera med en valfri riktning  $d$ : 
$$-\nabla f(\hat{x})^T d = \sum_i u_i \nabla g_i(\hat{x})^T d.$$

Om  $d$  är en **tillåten riktning**, så är  $\nabla g_i(\hat{x})^T d \leq 0$  för alla  $i$ .

Om  $u_i \geq 0$  för alla  $i$ , så blir 
$$\sum_i u_i \nabla g_i(\hat{x})^T d \leq 0,$$

vilket betyder att  $-\nabla f(\hat{x})^T d \leq 0$ , dvs.  $\nabla f(\hat{x})^T d \geq 0$ ,

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Vi har räknat ut  $u$  så att 
$$-\nabla f(\hat{x}) = \sum_i u_i \nabla g_i(\hat{x}).$$

Påstående: Om  $u_i \geq 0$  för alla  $i$ , så är alla förbättringsriktningar otillåtna.

Multiplitera med en valfri riktning  $d$ : 
$$-\nabla f(\hat{x})^T d = \sum_i u_i \nabla g_i(\hat{x})^T d.$$

Om  $d$  är en **tillåten riktning**, så är  $\nabla g_i(\hat{x})^T d \leq 0$  för alla  $i$ .

Om  $u_i \geq 0$  för alla  $i$ , så blir 
$$\sum_i u_i \nabla g_i(\hat{x})^T d \leq 0,$$

vilket betyder att  $-\nabla f(\hat{x})^T d \leq 0$ , dvs.  $\nabla f(\hat{x})^T d \geq 0$ ,

vilket visar att  $d$  **inte är en avtaganderiktning**.

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Vi har räknat ut  $u$  så att 
$$-\nabla f(\hat{x}) = \sum_i u_i \nabla g_i(\hat{x}).$$

Påstående: Om  $u_i \geq 0$  för alla  $i$ , så är alla förbättringsriktningar otillåtna.

Multiplitera med en valfri riktning  $d$ : 
$$-\nabla f(\hat{x})^T d = \sum_i u_i \nabla g_i(\hat{x})^T d.$$

Om  $d$  är en **tillåten riktning**, så är  $\nabla g_i(\hat{x})^T d \leq 0$  för alla  $i$ .

Om  $u_i \geq 0$  för alla  $i$ , så blir 
$$\sum_i u_i \nabla g_i(\hat{x})^T d \leq 0,$$

vilket betyder att  $-\nabla f(\hat{x})^T d \leq 0$ , dvs.  $\nabla f(\hat{x})^T d \geq 0$ ,

vilket visar att  $d$  **inte är en avtaganderiktning**.

Å andra sidan, om något  $u_i < 0$ , så gäller inte detta.

## Olinjär optimering med bivillkor: Optimalitetsvillkor

Vi har räknat ut  $u$  så att 
$$-\nabla f(\hat{x}) = \sum_i u_i \nabla g_i(\hat{x}).$$

Påstående: Om  $u_i \geq 0$  för alla  $i$ , så är alla förbättringsriktningar otillåtna.

Multiplitera med en valfri riktning  $d$ : 
$$-\nabla f(\hat{x})^T d = \sum_i u_i \nabla g_i(\hat{x})^T d.$$

Om  $d$  är en **tillåten riktning**, så är  $\nabla g_i(\hat{x})^T d \leq 0$  för alla  $i$ .

Om  $u_i \geq 0$  för alla  $i$ , så blir 
$$\sum_i u_i \nabla g_i(\hat{x})^T d \leq 0,$$

vilket betyder att  $-\nabla f(\hat{x})^T d \leq 0$ , dvs.  $\nabla f(\hat{x})^T d \geq 0$ ,

vilket visar att  $d$  **inte är en avtaganderiktning**.

Å andra sidan, om något  $u_i < 0$ , så gäller inte detta.

Så för säker optimalitet krävs: **KKT4**:  $u_i \geq 0$  för alla  $i$ .

## Olinjär optimering med bivillkor: Optimalitetsvillkor

KKT4 tillsammans med KKT3, KKT2 och KKT1 bevisar att ingen tillåten förbättringsriktning finns i punkten  $\hat{x}$ ,

## Olinjär optimering med bivillkor: Optimalitetsvillkor

KKT4 tillsammans med KKT3, KKT2 och KKT1 bevisar att ingen tillåten förbättringsriktning finns i punkten  $\hat{x}$ , så  $\hat{x}$  kan vara optimal.

## Olinjär optimering med bivillkor: Optimalitetsvillkor

KKT4 tillsammans med KKT3, KKT2 och KKT1 bevisar att ingen tillåten förbättringsriktning finns i punkten  $\hat{x}$ , så  $\hat{x}$  kan vara optimal.

En sådan punkt kallas KKT-punkt.



## Olinjär optimering med bivillkor: Optimalitetsvillkor

KKT4 tillsammans med KKT3, KKT2 och KKT1 bevisar att ingen tillåten förbättringsriktning finns i punkten  $\hat{x}$ , så  $\hat{x}$  kan vara optimal.

En sådan punkt kallas KKT-punkt.

För att en KKT-punkt säkert ska vara optimal, krävs **konvexitet**.

## Olinjär optimering med bivillkor: Optimalitetsvillkor

KKT4 tillsammans med KKT3, KKT2 och KKT1 bevisar att ingen tillåten förbättringsriktning finns i punkten  $\hat{x}$ , så  $\hat{x}$  kan vara optimal.

En sådan punkt kallas KKT-punkt.

För att en KKT-punkt säkert ska vara optimal, krävs **konvexitet**.

Annars kan det vara ett lokalt optimum eller en sadelpunkt (eller någon annan konstig punkt).

## Olinjär optimering med bivillkor: Optimalitetsvillkor

KKT4 tillsammans med KKT3, KKT2 och KKT1 bevisar att ingen tillåten förbättringsriktning finns i punkten  $\hat{x}$ , så  $\hat{x}$  kan vara optimal.

En sådan punkt kallas KKT-punkt.

För att en KKT-punkt säkert ska vara optimal, krävs **konvexitet**.

Annars kan det vara ett lokalt optimum eller en sadelpunkt (eller någon annan konstig punkt).

Detta är ett sätt att kontrollera om en viss *given* punkt är optimal.

## Olinjär optimering med bivillkor: Optimalitetsvillkor

KKT4 tillsammans med KKT3, KKT2 och KKT1 bevisar att ingen tillåten förbättringsriktning finns i punkten  $\hat{x}$ , så  $\hat{x}$  kan vara optimal.

En sådan punkt kallas KKT-punkt.

För att en KKT-punkt säkert ska vara optimal, krävs **konvexitet**.

Annars kan det vara ett lokalt optimum eller en sadelpunkt (eller någon annan konstig punkt).

Detta är ett sätt att kontrollera om en viss *given* punkt är optimal.

Då är alla gradienter givna, och bara  $u$  behöver lösas ut.

## Olinjär optimering med bivillkor: Optimalitetsvillkor

KKT4 tillsammans med KKT3, KKT2 och KKT1 bevisar att ingen tillåten förbättringsriktning finns i punkten  $\hat{x}$ , så  $\hat{x}$  kan vara optimal.

En sådan punkt kallas KKT-punkt.

För att en KKT-punkt säkert ska vara optimal, krävs **konvexitet**.

Annars kan det vara ett lokalt optimum eller en sadelpunkt (eller någon annan konstig punkt).

Detta är ett sätt att kontrollera om en viss *given* punkt är optimal.

Då är alla gradienter givna, och bara  $u$  behöver lösas ut.

KKT3 ger då ett *linjärt* ekvationssystem.

## Olinjär optimering med bivillkor: Optimalitetsvillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

# Olinjär optimering med bivillkor: Optimalitetsvillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

**Karush-Kuhn-Tuckervillkoren:**

## Olinjär optimering med bivillkor: Optimalitetsvillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

### Karush-Kuhn-Tuckervillkoren:

Punkten  $\hat{x}$  är en KKT-punkt om följande är uppfyllt.



## Olinjär optimering med bivillkor: Optimalitetsvillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

### Karush-Kuhn-Tuckervillkoren:

Punkten  $\hat{x}$  är en KKT-punkt om följande är uppfyllt.

**KKT1:**  $g_i(\hat{x}) \leq 0$  för alla  $i$ .

## Olinjär optimering med bivillkor: Optimalitetsvillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

### Karush-Kuhn-Tuckervillkoren:

Punkten  $\hat{x}$  är en KKT-punkt om följande är uppfyllt.

**KKT1:**  $g_i(\hat{x}) \leq 0$  för alla  $i$ . (Tillåtenhet)

# Olinjär optimering med bivillkor: Optimalitetsvillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

## Karush-Kuhn-Tuckervillkoren:

Punkten  $\hat{x}$  är en KKT-punkt om följande är uppfyllt.

**KKT1:**  $g_i(\hat{x}) \leq 0$  för alla  $i$ . (Tillåtenhet)

**KKT2:**  $u_i g_i(\hat{x}) = 0$  för alla  $i$ .

# Olinjär optimering med bivillkor: Optimalitetsvillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

## Karush-Kuhn-Tuckervillkoren:

Punkten  $\hat{x}$  är en KKT-punkt om följande är uppfyllt.

**KKT1:**  $g_i(\hat{x}) \leq 0$  för alla  $i$ . (Tillåtenhet)

**KKT2:**  $u_i g_i(\hat{x}) = 0$  för alla  $i$ . (Plocka bort inaktiva bivillkor)

# Olinjär optimering med bivillkor: Optimalitetsvillkor

$$\begin{aligned} \min \quad & f(x) \\ \text{då} \quad & g_i(x) \leq 0 \quad i = 1, \dots, m \end{aligned}$$

## Karush-Kuhn-Tuckervillkoren:

Punkten  $\hat{x}$  är en KKT-punkt om följande är uppfyllt.

**KKT1:**  $g_i(\hat{x}) \leq 0$  för alla  $i$ . (Tillåtenhet)

**KKT2:**  $u_i g_i(\hat{x}) = 0$  för alla  $i$ . (Plocka bort inaktiva bivillkor)

**KKT3:**  $\nabla f(\hat{x}) + \sum_{i=1}^m u_i \nabla g_i(\hat{x}) = 0$ .

# Olinjär optimering med bivillkor: Optimalitetsvillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

## Karush-Kuhn-Tuckervillkoren:

Punkten  $\hat{x}$  är en KKT-punkt om följande är uppfyllt.

**KKT1:**  $g_i(\hat{x}) \leq 0$  för alla  $i$ . (Tillåtenhet)

**KKT2:**  $u_i g_i(\hat{x}) = 0$  för alla  $i$ . (Plocka bort inaktiva bivillkor)

**KKT3:**  $\nabla f(\hat{x}) + \sum_{i=1}^m u_i \nabla g_i(\hat{x}) = 0$ . (Projektion)

# Olinjär optimering med bivillkor: Optimalitetsvillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

## Karush-Kuhn-Tuckervillkoren:

Punkten  $\hat{x}$  är en KKT-punkt om följande är uppfyllt.

**KKT1:**  $g_i(\hat{x}) \leq 0$  för alla  $i$ . (Tillåtenhet)

**KKT2:**  $u_i g_i(\hat{x}) = 0$  för alla  $i$ . (Plocka bort inaktiva bivillkor)

**KKT3:**  $\nabla f(\hat{x}) + \sum_{i=1}^m u_i \nabla g_i(\hat{x}) = 0$ . (Projektion)

**KKT4:**  $u_i \geq 0$  för alla  $i$ .

# Olinjär optimering med bivillkor: Optimalitetsvillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

## Karush-Kuhn-Tuckervillkoren:

Punkten  $\hat{x}$  är en KKT-punkt om följande är uppfyllt.

**KKT1:**  $g_i(\hat{x}) \leq 0$  för alla  $i$ . (Tillåtenhet)

**KKT2:**  $u_i g_i(\hat{x}) = 0$  för alla  $i$ . (Plocka bort inaktiva bivillkor)

**KKT3:**  $\nabla f(\hat{x}) + \sum_{i=1}^m u_i \nabla g_i(\hat{x}) = 0$ . (Projektion)

**KKT4:**  $u_i \geq 0$  för alla  $i$ . (Kolla riktning)



# Olinjär optimering med bivillkor: Optimalitetsvillkor

$$\begin{array}{ll} \min & f(x) \\ \text{då} & g_i(x) \leq 0 \quad i = 1, \dots, m \end{array}$$

## Karush-Kuhn-Tuckervillkoren:

Punkten  $\hat{x}$  är en KKT-punkt om följande är uppfyllt.

**KKT1:**  $g_i(\hat{x}) \leq 0$  för alla  $i$ . (Tillåtenhet)

**KKT2:**  $u_i g_i(\hat{x}) = 0$  för alla  $i$ . (Plocka bort inaktiva bivillkor)

**KKT3:**  $\nabla f(\hat{x}) + \sum_{i=1}^m u_i \nabla g_i(\hat{x}) = 0$ . (Projektion)

**KKT4:**  $u_i \geq 0$  för alla  $i$ . (Kolla riktning)

KKT3 är ett linjärt ekvationssystem i  $u$  som kan lösas metodiskt.

## Tolkning av KKT-villkoren

KKT1 tar bort otillåtna punkter.

## Tolkning av KKT-villkoren

KKT1 tar bort otillåtna punkter.

KKT2 tar bort inaktiva bivillkor.

## Tolkning av KKT-villkoren

KKT1 tar bort otillåtna punkter.

KKT2 tar bort inaktiva bivillkor.

KKT3 projicerar önskeriktningen  $-\nabla f(x)$  på aktiva bivillkors normaler  $\nabla g_i(x)$ .

## Tolkning av KKT-villkoren

KKT1 tar bort otillåtna punkter.

KKT2 tar bort inaktiva bivillkor.

KKT3 projicerar önskeriktningen  $-\nabla f(x)$  på aktiva bivillkors normaler  $\nabla g_i(x)$ .

KKT4 kontrollerar att önskeriktningen  $-\nabla f(x)$  ligger i konen som spänns upp av aktiva bivillkors normaler  $\nabla g_i(x)$ ,

## Tolkning av KKT-villkoren

KKT1 tar bort otillåtna punkter.

KKT2 tar bort inaktiva bivillkor.

KKT3 projicerar önskeriktningen  $-\nabla f(x)$  på aktiva bivillkors normaler  $\nabla g_i(x)$ .

KKT4 kontrollerar att önskeriktningen  $-\nabla f(x)$  ligger i konen som spänns upp av aktiva bivillkors normaler  $\nabla g_i(x)$ ,

dvs. att alla förbättringsriktningar är otillåtna.

## Tolkning av KKT-villkoren

KKT1 tar bort otillåtna punkter.

KKT2 tar bort inaktiva bivillkor.

KKT3 projicerar önskeriktningen  $-\nabla f(x)$  på aktiva bivillkors normaler  $\nabla g_i(x)$ .

KKT4 kontrollerar att önskeriktningen  $-\nabla f(x)$  ligger i konen som spänns upp av aktiva bivillkors normaler  $\nabla g_i(x)$ ,

dvs. att alla förbättringsriktningar är otillåtna.

Hur gör man med likhetsbivillkor?

## Tolkning av KKT-villkoren

KKT1 tar bort otillåtna punkter.

KKT2 tar bort inaktiva bivillkor.

KKT3 projicerar önskeriktningen  $-\nabla f(x)$  på aktiva bivillkors normaler  $\nabla g_i(x)$ .

KKT4 kontrollerar att önskeriktningen  $-\nabla f(x)$  ligger i konen som spänns upp av aktiva bivillkors normaler  $\nabla g_i(x)$ ,

dvs. att alla förbättringsriktningar är otillåtna.

Hur gör man med likhetsbivillkor?

De är alltid aktiva, och  $u_i < 0$  är tillåtet, så KKT2 och KKT4 behövs inte för dessa bivillkor.



## Vårt tidigare exempel

$$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2 \text{ då } x_1 + x_2 \leq 1, x_2 \leq 0.4.$$

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

$$g_1(\hat{x}) = 0.$$

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

$g_1(\hat{x}) = 0$ . Tillåtet.

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt.

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_1$  behöver inte vara noll.



## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_1$  behöver inte vara noll.

$g_2(\hat{x}) = 0$ .

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_1$  behöver inte vara noll.

$g_2(\hat{x}) = 0$ . Tillåtet.

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_1$  behöver inte vara noll.

$g_2(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt.

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_1$  behöver inte vara noll.

$g_2(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_2$  behöver inte vara noll.

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_1$  behöver inte vara noll.

$g_2(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_2$  behöver inte vara noll.

Punkten är tillåten.

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_1$  behöver inte vara noll.

$g_2(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_2$  behöver inte vara noll.

Punkten är tillåten. KKT2: Inget av  $u_1$  eller  $u_2$  måste vara noll.

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_1$  behöver inte vara noll.

$g_2(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_2$  behöver inte vara noll.

Punkten är tillåten. KKT2: Inget av  $u_1$  eller  $u_2$  måste vara noll.

$$\text{KKT3: } \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_1$  behöver inte vara noll.

$g_2(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_2$  behöver inte vara noll.

Punkten är tillåten. KKT2: Inget av  $u_1$  eller  $u_2$  måste vara noll.

$$\text{KKT3: } \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Dvs.  $u_1 = 0.2$ ,  $u_1 + u_2 = 0.8$ ,



## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_1$  behöver inte vara noll.

$g_2(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_2$  behöver inte vara noll.

Punkten är tillåten. KKT2: Inget av  $u_1$  eller  $u_2$  måste vara noll.

$$\text{KKT3: } \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Dvs.  $u_1 = 0.2$ ,  $u_1 + u_2 = 0.8$ , vilket ger  $u_1 = 0.2$ ,  $u_2 = 0.6$ .

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_1$  behöver inte vara noll.

$g_2(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_2$  behöver inte vara noll.

Punkten är tillåten. KKT2: Inget av  $u_1$  eller  $u_2$  måste vara noll.

$$\text{KKT3: } \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Dvs.  $u_1 = 0.2$ ,  $u_1 + u_2 = 0.8$ , vilket ger  $u_1 = 0.2$ ,  $u_2 = 0.6$ .

KKT4 uppfyllt,  $u_1 \geq 0$  och  $u_2 \geq 0$ .

## Vårt tidigare exempel

$\min x_1^2 + x_2^2 - x_1x_2 - x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_2 \leq 0.4$ .

Är  $\hat{x}_1 = 0.6$ ,  $\hat{x}_2 = 0.4$  optimal?

$$\nabla f(\hat{x}) = \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

KKT1: Är punkten tillåten?

$g_1(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_1$  behöver inte vara noll.

$g_2(\hat{x}) = 0$ . Tillåtet. Bivillkoret aktivt. KKT2:  $u_2$  behöver inte vara noll.

Punkten är tillåten. KKT2: Inget av  $u_1$  eller  $u_2$  måste vara noll.

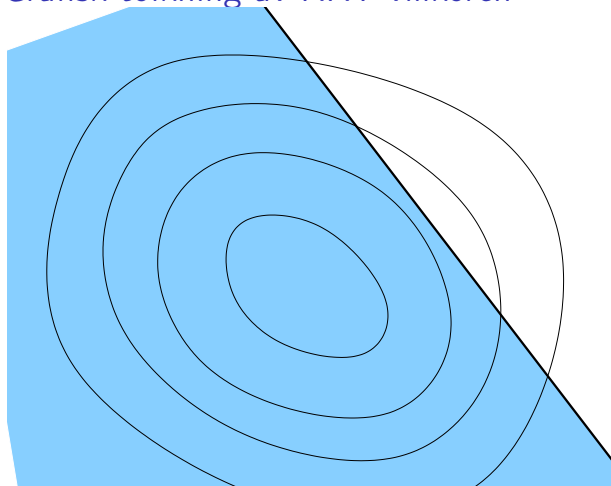
$$\text{KKT3: } \begin{pmatrix} -0.2 \\ -0.8 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Dvs.  $u_1 = 0.2$ ,  $u_1 + u_2 = 0.8$ , vilket ger  $u_1 = 0.2$ ,  $u_2 = 0.6$ .

KKT4 uppfyllt,  $u_1 \geq 0$  och  $u_2 \geq 0$ .

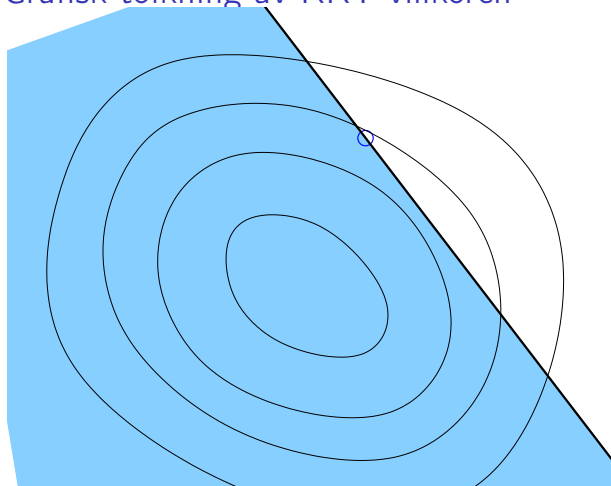
Punkten är optimal, ty den är en KKT-punkt och problemet är konvext.

## Grafisk tolkning av KKT-villkoren



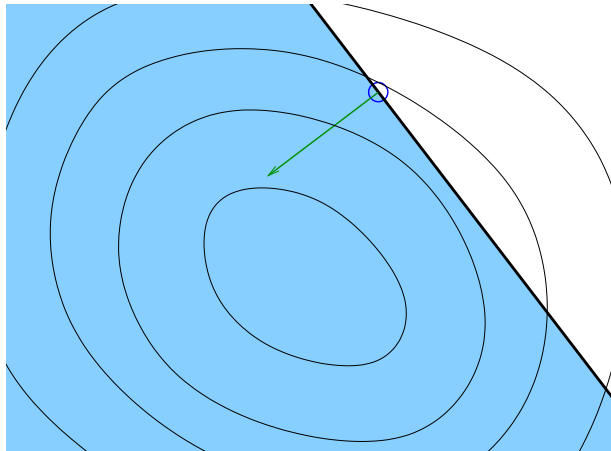
Tillåtet område och nivåkurvor.

## Grafisk tolkning av KKT-villkoren



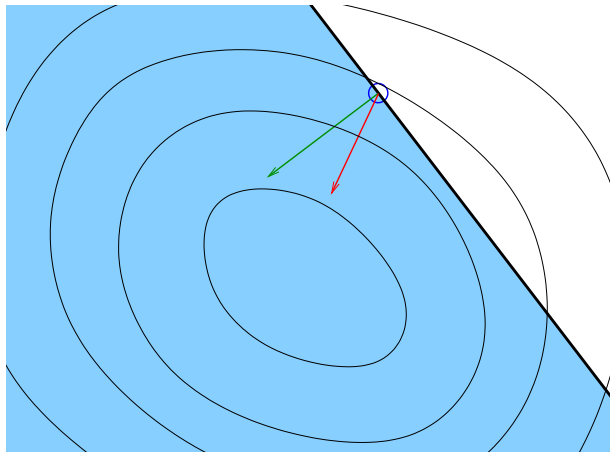
Är denna punkt optimal?

## Grafisk tolkning av KKT-villkoren



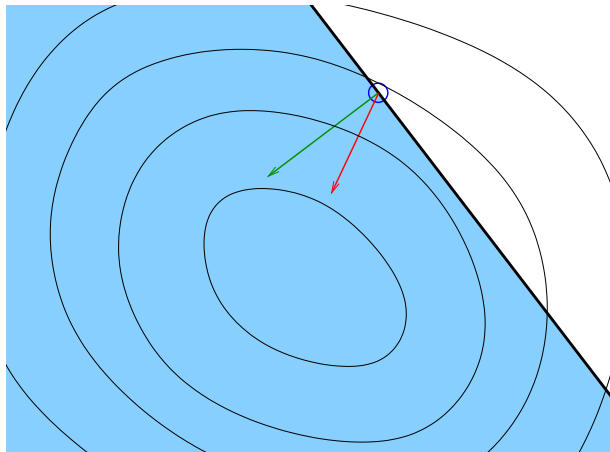
Utåtriktad bivillkorsnormal ( $\nabla g(x)$ ).

## Grafisk tolkning av KKT-villkoren



Önskeriktning ( $-\nabla f(x)$ ).

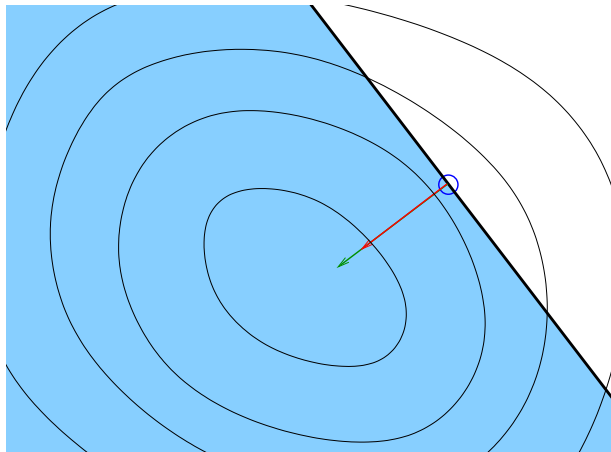
## Grafisk tolkning av KKT-villkoren



Projektion ej möjlig. Ej KKT-punkt. Ej optimum.

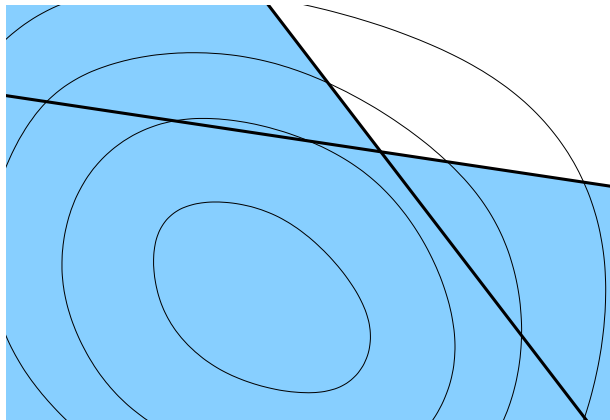


## Grafisk tolkning av KKT-villkoren



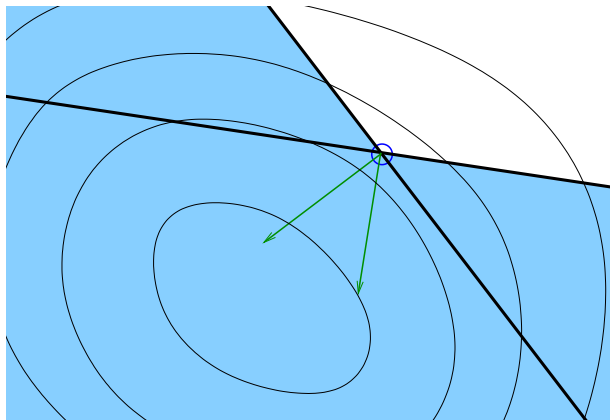
En punkt som är optimal.

## Grafisk tolkning av KKT-villkoren



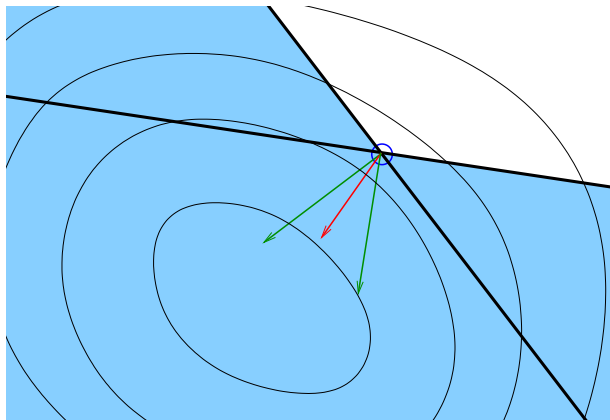
Två bivillkor.

## Grafisk tolkning av KKT-villkoren



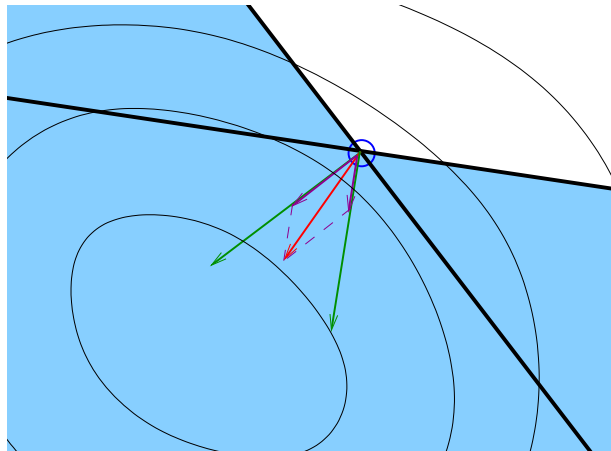
Bivillkorsgradienter.

## Grafisk tolkning av KKT-villkoren



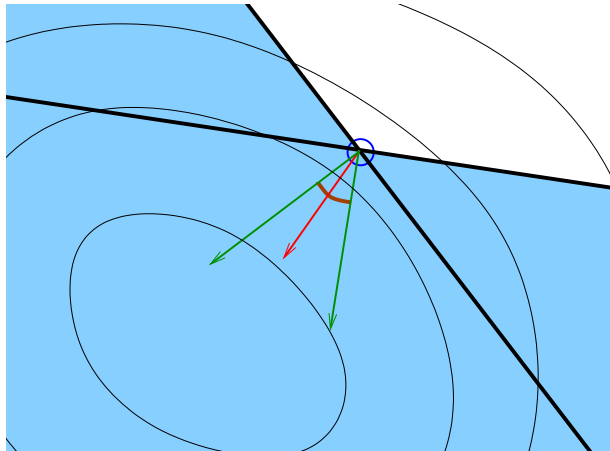
Önskeriktning.

## Grafisk tolkning av KKT-villkoren



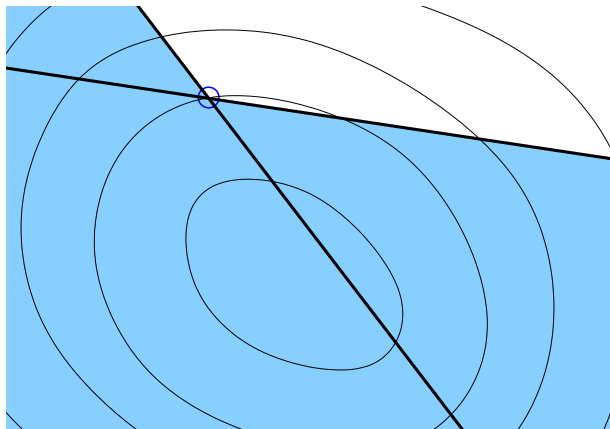
Projektion.  $u \geq 0$ . Optimum.

## Grafisk tolkning av KKT-villkoren



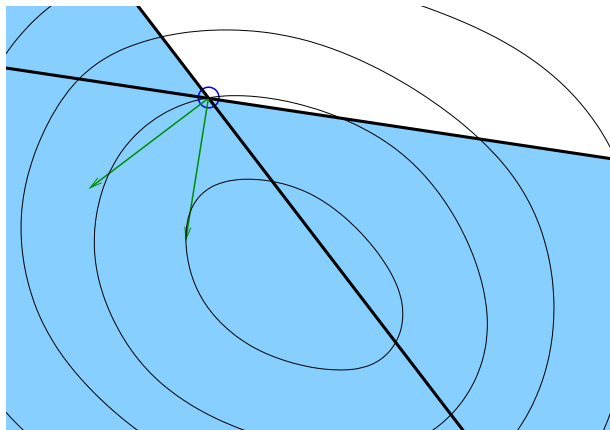
Önskeriktningen ligger i konen av aktiva bivillkorsgradients.

## Grafisk tolkning av KKT-villkoren



Ett bivillkor flyttat.

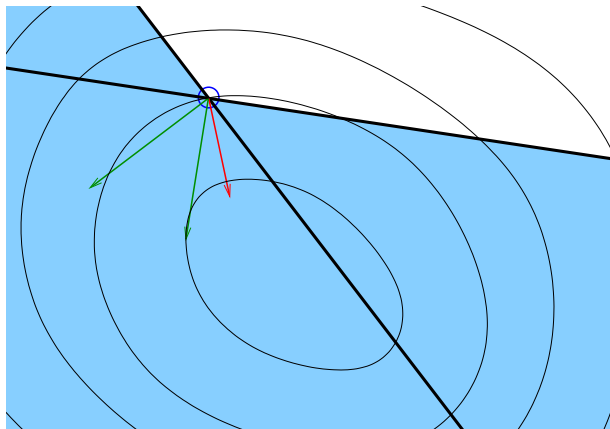
## Grafisk tolkning av KKT-villkoren



Bivillkorsgradienter.

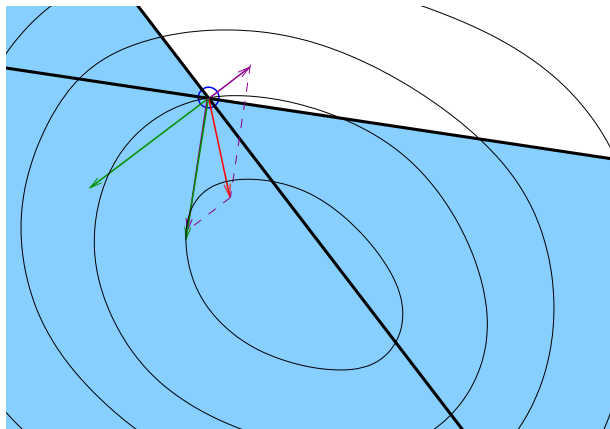


## Grafisk tolkning av KKT-villkoren



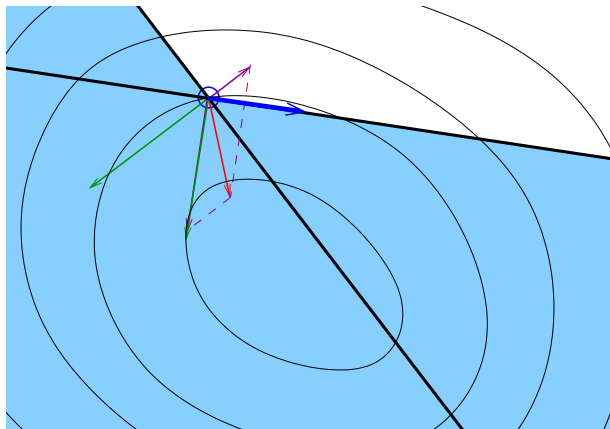
Önskeriktning.

## Grafisk tolkning av KKT-villkoren



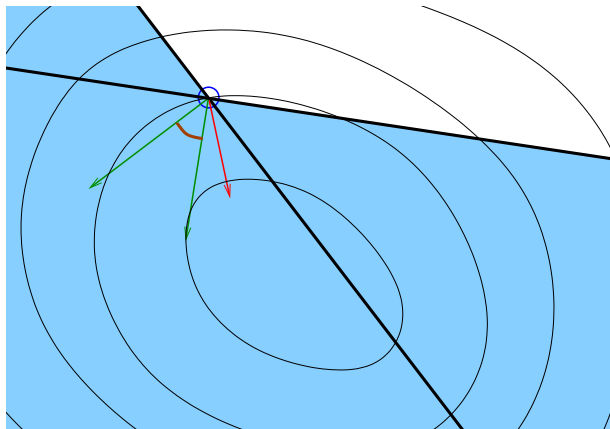
Projektion. Ett  $u < 0$ . Ej optimum.

## Grafisk tolkning av KKT-villkoren



En tillåten förbättringsriktning finns.

## Grafisk tolkning av KKT-villkoren



Önskeriktningen ligger *inte* i konen av aktiva bivillkorsgradients.

# KKT-villkoren och optimalitet

En KKT-punkt kan vara optimum.

# KKT-villkoren och optimalitet

En KKT-punkt kan vara optimum.

Om problemet är *konvext*, så är en KKT-punkt ett *globalt optimum*.

# KKT-villkoren och optimalitet

En KKT-punkt kan vara optimum.

Om problemet är *konvext*, så är en KKT-punkt ett *globalt optimum*.

Om problemet inte är konvext kan det finnas KKT-punkter som inte är optimala.

# KKT-villkoren och optimalitet

En KKT-punkt kan vara optimum.

Om problemet är *konvext*, så är en KKT-punkt ett *globalt optimum*.

Om problemet inte är konvext kan det finnas KKT-punkter som inte är optimala. Vissa KKT-punkter kan vara lokala optima.



# KKT-villkoren och optimalitet

En KKT-punkt kan vara optimum.

Om problemet är *konvext*, så är en KKT-punkt ett *globalt optimum*.

Om problemet inte är konvext kan det finnas KKT-punkter som inte är optimala. Vissa KKT-punkter kan vara lokala optima.

Det finns vissa krav, kallade CQ (Constraint Qualification), för att optimum ska vara en KKT-punkt. Om inget CQ är uppfyllt, kanske det inte finns någon KKT-punkt.

## KKT-villkoren och optimalitet

En KKT-punkt kan vara optimum.

Om problemet är *konvext*, så är en KKT-punkt ett *globalt optimum*.

Om problemet inte är konvext kan det finnas KKT-punkter som inte är optimala. Vissa KKT-punkter kan vara lokala optima.

Det finns vissa krav, kallade CQ (Constraint Qualification), för att optimum ska vara en KKT-punkt. Om inget CQ är uppfyllt, kanske det inte finns någon KKT-punkt. (CQ kan i denna kurs antas vara uppfyllda.)

# KKT-villkoren och optimalitet

En KKT-punkt kan vara optimum.

Om problemet är *konvext*, så är en KKT-punkt ett *globalt optimum*.

Om problemet inte är konvext kan det finnas KKT-punkter som inte är optimala. Vissa KKT-punkter kan vara lokala optima.

Det finns vissa krav, kallade CQ (Constraint Qualification), för att optimum ska vara en KKT-punkt. Om inget CQ är uppfyllt, kanske det inte finns någon KKT-punkt. (CQ kan i denna kurs antas vara uppfyllda.)

Om något CQ är uppfyllt, och problemet är konvext, är alltså optimum och bara optimum är en KKT-punkt.

## KKT-villkoren: Exempel

Exempel:  $\min f(x) = x_1^2 + 2x_2^2$  då  $g(x) = x_1 + x_2 + 1 \leq 0$ .

## KKT-villkoren: Exempel

Exempel:  $\min f(x) = x_1^2 + 2x_2^2$  då  $g(x) = x_1 + x_2 + 1 \leq 0$ .

Är  $\hat{x}_1 = -0.5$  och  $\hat{x}_2 = -0.5$  optimal?

## KKT-villkoren: Exempel

Exempel:  $\min f(x) = x_1^2 + 2x_2^2$  då  $g(x) = x_1 + x_2 + 1 \leq 0$ .

Är  $\hat{x}_1 = -0.5$  och  $\hat{x}_2 = -0.5$  optimal?

KKT1: Punkten är tillåten.

## KKT-villkoren: Exempel

Exempel:  $\min f(x) = x_1^2 + 2x_2^2$  då  $g(x) = x_1 + x_2 + 1 \leq 0$ .

Är  $\hat{x}_1 = -0.5$  och  $\hat{x}_2 = -0.5$  optimal?

KKT1: Punkten är tillåten.

KKT2: Bivillkoret är aktivt.

## KKT-villkoren: Exempel

Exempel:  $\min f(x) = x_1^2 + 2x_2^2$  då  $g(x) = x_1 + x_2 + 1 \leq 0$ .

Är  $\hat{x}_1 = -0.5$  och  $\hat{x}_2 = -0.5$  optimal?

KKT1: Punkten är tillåten.

KKT2: Bivillkoret är aktivt. (Så  $u$  behöver inte vara 0.)



## KKT-villkoren: Exempel

Exempel:  $\min f(x) = x_1^2 + 2x_2^2$  då  $g(x) = x_1 + x_2 + 1 \leq 0$ .

Är  $\hat{x}_1 = -0.5$  och  $\hat{x}_2 = -0.5$  optimal?

KKT1: Punkten är tillåten.

KKT2: Bivillkoret är aktivt. (Så  $u$  behöver inte vara 0.)

KKT3:  $\nabla f(\hat{x}) = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$ ,

## KKT-villkoren: Exempel

Exempel:  $\min f(x) = x_1^2 + 2x_2^2$  då  $g(x) = x_1 + x_2 + 1 \leq 0$ .

Är  $\hat{x}_1 = -0.5$  och  $\hat{x}_2 = -0.5$  optimal?

KKT1: Punkten är tillåten.

KKT2: Bivillkoret är aktivt. (Så  $u$  behöver inte vara 0.)

KKT3:  $\nabla f(\hat{x}) = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$ , och  $\nabla g(\hat{x}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ .

## KKT-villkoren: Exempel

Exempel:  $\min f(x) = x_1^2 + 2x_2^2$  då  $g(x) = x_1 + x_2 + 1 \leq 0$ .

Är  $\hat{x}_1 = -0.5$  och  $\hat{x}_2 = -0.5$  optimal?

KKT1: Punkten är tillåten.

KKT2: Bivillkoret är aktivt. (Så  $u$  behöver inte vara 0.)

KKT3:  $\nabla f(\hat{x}) = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$ , och  $\nabla g(\hat{x}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ .

KKT3 blir  $\begin{pmatrix} -1 \\ -2 \end{pmatrix} + u \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ,

## KKT-villkoren: Exempel

Exempel:  $\min f(x) = x_1^2 + 2x_2^2$  då  $g(x) = x_1 + x_2 + 1 \leq 0$ .

Är  $\hat{x}_1 = -0.5$  och  $\hat{x}_2 = -0.5$  optimal?

KKT1: Punkten är tillåten.

KKT2: Bivillkoret är aktivt. (Så  $u$  behöver inte vara 0.)

KKT3:  $\nabla f(\hat{x}) = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$ , och  $\nabla g(\hat{x}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ .

KKT3 blir  $\begin{pmatrix} -1 \\ -2 \end{pmatrix} + u \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ,

vilket betyder  $u = 1$  och  $u = 2$ .

## KKT-villkoren: Exempel

Exempel:  $\min f(x) = x_1^2 + 2x_2^2$  då  $g(x) = x_1 + x_2 + 1 \leq 0$ .

Är  $\hat{x}_1 = -0.5$  och  $\hat{x}_2 = -0.5$  optimal?

KKT1: Punkten är tillåten.

KKT2: Bivillkoret är aktivt. (Så  $u$  behöver inte vara 0.)

KKT3:  $\nabla f(\hat{x}) = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$ , och  $\nabla g(\hat{x}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ .

KKT3 blir  $\begin{pmatrix} -1 \\ -2 \end{pmatrix} + u \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ,

vilket betyder  $u = 1$  och  $u = 2$ . Ej lösbart.

## KKT-villkoren: Exempel

Exempel:  $\min f(x) = x_1^2 + 2x_2^2$  då  $g(x) = x_1 + x_2 + 1 \leq 0$ .

Är  $\hat{x}_1 = -0.5$  och  $\hat{x}_2 = -0.5$  optimal?

KKT1: Punkten är tillåten.

KKT2: Bivillkoret är aktivt. (Så  $u$  behöver inte vara 0.)

KKT3:  $\nabla f(\hat{x}) = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$ , och  $\nabla g(\hat{x}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ .

KKT3 blir  $\begin{pmatrix} -1 \\ -2 \end{pmatrix} + u \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ,

vilket betyder  $u = 1$  och  $u = 2$ . Ej lösbart.

$\hat{x}$  är alltså ingen KKT-punkt.

## KKT-villkoren: Exempel

Exempel:  $\min x_1^2 + 2x_2^2$  då  $x_1 + x_2 \geq 1$ ,  $x_1 \geq 0$ ,  $x_2 \geq 0$ .

## KKT-villkoren: Exempel

Exempel:  $\min x_1^2 + 2x_2^2$  då  $x_1 + x_2 \geq 1$ ,  $x_1 \geq 0$ ,  $x_2 \geq 0$ .

Skriv bivillkoren som

$$g_1(x) = -x_1 - x_2 + 1 \leq 0, \quad g_2(x) = -x_1 \leq 0, \quad g_3(x) = -x_2 \leq 0.$$



## KKT-villkoren: Exempel

Exempel:  $\min x_1^2 + 2x_2^2$  då  $x_1 + x_2 \geq 1$ ,  $x_1 \geq 0$ ,  $x_2 \geq 0$ .

Skriv bivillkoren som

$$g_1(x) = -x_1 - x_2 + 1 \leq 0, \quad g_2(x) = -x_1 \leq 0, \quad g_3(x) = -x_2 \leq 0.$$

$$\text{Gradienter: } \nabla f(x) = \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} -1 \\ 0 \end{pmatrix},$$

$$\nabla g_3(x) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

## KKT-villkoren: Exempel

Exempel:  $\min x_1^2 + 2x_2^2$  då  $x_1 + x_2 \geq 1$ ,  $x_1 \geq 0$ ,  $x_2 \geq 0$ .

Skriv bivillkoren som

$$g_1(x) = -x_1 - x_2 + 1 \leq 0, \quad g_2(x) = -x_1 \leq 0, \quad g_3(x) = -x_2 \leq 0.$$

$$\text{Gradienter: } \nabla f(x) = \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} -1 \\ 0 \end{pmatrix},$$

$$\nabla g_3(x) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

$$\text{KKT3: } \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + u_3 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

## KKT-villkoren: Exempel

Exempel:  $\min x_1^2 + 2x_2^2$  då  $x_1 + x_2 \geq 1$ ,  $x_1 \geq 0$ ,  $x_2 \geq 0$ .

Skriv bivillkoren som

$$g_1(x) = -x_1 - x_2 + 1 \leq 0, \quad g_2(x) = -x_1 \leq 0, \quad g_3(x) = -x_2 \leq 0.$$

$$\text{Gradienter: } \nabla f(x) = \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} -1 \\ 0 \end{pmatrix},$$

$$\nabla g_3(x) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

$$\text{KKT3: } \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + u_3 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Sätt in de punkter som ska testas.

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 0$  optimal?

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 0$  optimal?

KKT1: Punkten är tillåten.

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 0$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ .

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 0$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 0$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = -1 < 0$ .



## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 0$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = -1 < 0$ . ( $u_2 = 0$ .)

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 0$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = -1 < 0$ . ( $u_2 = 0$ .)

KKT2:  $g_3(\hat{x}) = 0$ .

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 0$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = -1 < 0$ . ( $u_2 = 0$ .)

KKT2:  $g_3(\hat{x}) = 0$ . ( $u_3$  behöver inte vara 0.)

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 0$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = -1 < 0$ . ( $u_2 = 0$ .)

KKT2:  $g_3(\hat{x}) = 0$ . ( $u_3$  behöver inte vara 0.)

KKT3:  $\nabla f(\hat{x}) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ .

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 0$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = -1 < 0$ . ( $u_2 = 0$ .)

KKT2:  $g_3(\hat{x}) = 0$ . ( $u_3$  behöver inte vara 0.)

KKT3:  $\nabla f(\hat{x}) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ .

KKT3 blir  $\begin{pmatrix} 2 \\ 0 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} + u_3 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ,

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 0$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = -1 < 0$ . ( $u_2 = 0$ .)

KKT2:  $g_3(\hat{x}) = 0$ . ( $u_3$  behöver inte vara 0.)

$$\text{KKT3: } \nabla f(\hat{x}) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

$$\text{KKT3 blir } \begin{pmatrix} 2 \\ 0 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} + u_3 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\text{dvs. } 2 - u_1 = 0, \quad -u_1 - u_3 = 0,$$

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 0$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = -1 < 0$ . ( $u_2 = 0$ .)

KKT2:  $g_3(\hat{x}) = 0$ . ( $u_3$  behöver inte vara 0.)

$$\text{KKT3: } \nabla f(\hat{x}) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

$$\text{KKT3 blir } \begin{pmatrix} 2 \\ 0 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} + u_3 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $2 - u_1 = 0$ ,  $-u_1 - u_3 = 0$ , vilket har lösningen  $u_1 = 2$  och  $u_3 = -2$ .

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 0$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = -1 < 0$ . ( $u_2 = 0$ .)

KKT2:  $g_3(\hat{x}) = 0$ . ( $u_3$  behöver inte vara 0.)

$$\text{KKT3: } \nabla f(\hat{x}) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

$$\text{KKT3 blir } \begin{pmatrix} 2 \\ 0 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} + u_3 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $2 - u_1 = 0$ ,  $-u_1 - u_3 = 0$ , vilket har lösningen  $u_1 = 2$  och  $u_3 = -2$ .

$\hat{x}$  är ingen KKT-punkt, ty  $u_3 < 0$ .



## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 0$  och  $\hat{x}_2 = 1$  optimal?

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 0$  och  $\hat{x}_2 = 1$  optimal?

KKT1: Punkten är tillåten.

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 0$  och  $\hat{x}_2 = 1$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ .

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 0$  och  $\hat{x}_2 = 1$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 0$  och  $\hat{x}_2 = 1$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = 0$ .

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 0$  och  $\hat{x}_2 = 1$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = 0$ . ( $u_2$  behöver inte vara 0.)

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 0$  och  $\hat{x}_2 = 1$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = 0$ . ( $u_2$  behöver inte vara 0.)

KKT2:  $g_3(\hat{x}) = -1 < 0$ .

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 0$  och  $\hat{x}_2 = 1$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = 0$ . ( $u_2$  behöver inte vara 0.)

KKT2:  $g_3(\hat{x}) = -1 < 0$ . ( $u_3 = 0$ .)



## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 0$  och  $\hat{x}_2 = 1$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = 0$ . ( $u_2$  behöver inte vara 0.)

KKT2:  $g_3(\hat{x}) = -1 < 0$ . ( $u_3 = 0$ .)

KKT3:  $\nabla f(\hat{x}) = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$ .

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 0$  och  $\hat{x}_2 = 1$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = 0$ . ( $u_2$  behöver inte vara 0.)

KKT2:  $g_3(\hat{x}) = -1 < 0$ . ( $u_3 = 0$ .)

KKT3:  $\nabla f(\hat{x}) = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$ .

KKT3 blir  $\begin{pmatrix} 0 \\ 4 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ,

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 0$  och  $\hat{x}_2 = 1$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = 0$ . ( $u_2$  behöver inte vara 0.)

KKT2:  $g_3(\hat{x}) = -1 < 0$ . ( $u_3 = 0$ .)

$$\text{KKT3: } \nabla f(\hat{x}) = \begin{pmatrix} 0 \\ 4 \end{pmatrix}.$$

$$\text{KKT3 blir } \begin{pmatrix} 0 \\ 4 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\text{dvs. } -u_1 - u_2 = 0, \quad 4 - u_1 = 0,$$

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 0$  och  $\hat{x}_2 = 1$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = 0$ . ( $u_2$  behöver inte vara 0.)

KKT2:  $g_3(\hat{x}) = -1 < 0$ . ( $u_3 = 0$ .)

$$\text{KKT3: } \nabla f(\hat{x}) = \begin{pmatrix} 0 \\ 4 \end{pmatrix}.$$

$$\text{KKT3 blir } \begin{pmatrix} 0 \\ 4 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $-u_1 - u_2 = 0$ ,  $4 - u_1 = 0$ , vilket har lösningen  $u_1 = 4$  och  $u_2 = -4$ .

## KKT-villkoren: Exempel

Är  $\hat{x}_1 = 0$  och  $\hat{x}_2 = 1$  optimal?

KKT1: Punkten är tillåten.

KKT2:  $g_1(\hat{x}) = 0$ . ( $u_1$  behöver inte vara 0.)

KKT2:  $g_2(\hat{x}) = 0$ . ( $u_2$  behöver inte vara 0.)

KKT2:  $g_3(\hat{x}) = -1 < 0$ . ( $u_3 = 0$ .)

$$\text{KKT3: } \nabla f(\hat{x}) = \begin{pmatrix} 0 \\ 4 \end{pmatrix}.$$

$$\text{KKT3 blir } \begin{pmatrix} 0 \\ 4 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $-u_1 - u_2 = 0$ ,  $4 - u_1 = 0$ , vilket har lösningen  $u_1 = 4$  och  $u_2 = -4$ .

$\hat{x}$  är ingen KKT-punkt, ty  $u_2 < 0$ .

## KKT-villkoren: Exempel

Istället för att ange en punkt kan man anta vilka bivillkor som är aktiva.

## KKT-villkoren: Exempel

Istället för att ange en punkt kan man anta vilka bivillkor som är aktiva.  
(Detta fungerar bara ibland.)

## KKT-villkoren: Exempel

Istället för att ange en punkt kan man anta vilka bivillkor som är aktiva.  
(Detta fungerar bara ibland.)

Antag att  $x_1 + x_2 = 1$  är aktivt och att  $x_1 > 0$  och  $x_2 > 0$ .



## KKT-villkoren: Exempel

Istället för att ange en punkt kan man anta vilka bivillkor som är aktiva.  
(Detta fungerar bara ibland.)

Antag att  $x_1 + x_2 = 1$  är aktivt och att  $x_1 > 0$  och  $x_2 > 0$ .

Dvs.  $g_1(\hat{x}) = 0$ ,  $g_2(\hat{x}) < 0$  och  $g_3(\hat{x}) < 0$ .

## KKT-villkoren: Exempel

Istället för att ange en punkt kan man anta vilka bivillkor som är aktiva.  
(Detta fungerar bara ibland.)

Antag att  $x_1 + x_2 = 1$  är aktivt och att  $x_1 > 0$  och  $x_2 > 0$ .

Dvs.  $g_1(\hat{x}) = 0$ ,  $g_2(\hat{x}) < 0$  och  $g_3(\hat{x}) < 0$ .

KKT2 ger då  $u_2 = 0$  och  $u_3 = 0$ .

## KKT-villkoren: Exempel

Istället för att ange en punkt kan man anta vilka bivillkor som är aktiva.  
(Detta fungerar bara ibland.)

Antag att  $x_1 + x_2 = 1$  är aktivt och att  $x_1 > 0$  och  $x_2 > 0$ .

Dvs.  $g_1(\hat{x}) = 0$ ,  $g_2(\hat{x}) < 0$  och  $g_3(\hat{x}) < 0$ .

KKT2 ger då  $u_2 = 0$  och  $u_3 = 0$ .

KKT3 blir då 
$$\begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

## KKT-villkoren: Exempel

Istället för att ange en punkt kan man anta vilka bivillkor som är aktiva.  
(Detta fungerar bara ibland.)

Antag att  $x_1 + x_2 = 1$  är aktivt och att  $x_1 > 0$  och  $x_2 > 0$ .

Dvs.  $g_1(\hat{x}) = 0$ ,  $g_2(\hat{x}) < 0$  och  $g_3(\hat{x}) < 0$ .

KKT2 ger då  $u_2 = 0$  och  $u_3 = 0$ .

$$\text{KKT3 blir då } \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

vilket betyder  $2x_1 - u_1 = 0$  och  $4x_2 - u_1 = 0$ .

## KKT-villkoren: Exempel

Istället för att ange en punkt kan man anta vilka bivillkor som är aktiva.  
(Detta fungerar bara ibland.)

Antag att  $x_1 + x_2 = 1$  är aktivt och att  $x_1 > 0$  och  $x_2 > 0$ .

Dvs.  $g_1(\hat{x}) = 0$ ,  $g_2(\hat{x}) < 0$  och  $g_3(\hat{x}) < 0$ .

KKT2 ger då  $u_2 = 0$  och  $u_3 = 0$ .

$$\text{KKT3 blir då } \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

vilket betyder  $2x_1 - u_1 = 0$  och  $4x_2 - u_1 = 0$ .

Lös ut  $x$  i  $u$ :  $x_1 = u_1/2$  och  $x_2 = u_1/4$ .

## KKT-villkoren: Exempel

Istället för att ange en punkt kan man anta vilka bivillkor som är aktiva.  
(Detta fungerar bara ibland.)

Antag att  $x_1 + x_2 = 1$  är aktivt och att  $x_1 > 0$  och  $x_2 > 0$ .

Dvs.  $g_1(\hat{x}) = 0$ ,  $g_2(\hat{x}) < 0$  och  $g_3(\hat{x}) < 0$ .

KKT2 ger då  $u_2 = 0$  och  $u_3 = 0$ .

$$\text{KKT3 blir då } \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

vilket betyder  $2x_1 - u_1 = 0$  och  $4x_2 - u_1 = 0$ .

Lös ut  $x$  i  $u$ :  $x_1 = u_1/2$  och  $x_2 = u_1/4$ .

Antagandet  $x_1 + x_2 = 1$  ger då  $x_1 + x_2 = u_1/2 + u_1/4 = 1$ ,

## KKT-villkoren: Exempel

Istället för att ange en punkt kan man anta vilka bivillkor som är aktiva.  
(Detta fungerar bara ibland.)

Antag att  $x_1 + x_2 = 1$  är aktivt och att  $x_1 > 0$  och  $x_2 > 0$ .

Dvs.  $g_1(\hat{x}) = 0$ ,  $g_2(\hat{x}) < 0$  och  $g_3(\hat{x}) < 0$ .

KKT2 ger då  $u_2 = 0$  och  $u_3 = 0$ .

$$\text{KKT3 blir då } \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

vilket betyder  $2x_1 - u_1 = 0$  och  $4x_2 - u_1 = 0$ .

Lös ut  $x$  i  $u$ :  $x_1 = u_1/2$  och  $x_2 = u_1/4$ .

Antagandet  $x_1 + x_2 = 1$  ger då  $x_1 + x_2 = u_1/2 + u_1/4 = 1$ ,  
vilket ger  $u_1 = 4/3$ ,

## KKT-villkoren: Exempel

Istället för att ange en punkt kan man anta vilka bivillkor som är aktiva.  
(Detta fungerar bara ibland.)

Antag att  $x_1 + x_2 = 1$  är aktivt och att  $x_1 > 0$  och  $x_2 > 0$ .

Dvs.  $g_1(\hat{x}) = 0$ ,  $g_2(\hat{x}) < 0$  och  $g_3(\hat{x}) < 0$ .

KKT2 ger då  $u_2 = 0$  och  $u_3 = 0$ .

$$\text{KKT3 blir då } \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

vilket betyder  $2x_1 - u_1 = 0$  och  $4x_2 - u_1 = 0$ .

Lös ut  $x$  i  $u$ :  $x_1 = u_1/2$  och  $x_2 = u_1/4$ .

Antagandet  $x_1 + x_2 = 1$  ger då  $x_1 + x_2 = u_1/2 + u_1/4 = 1$ ,  
vilket ger  $u_1 = 4/3$ , och  $x_1 = 2/3$  och  $x_2 = 1/3$ .



## KKT-villkoren: Exempel

Istället för att ange en punkt kan man anta vilka bivillkor som är aktiva.  
(Detta fungerar bara ibland.)

Antag att  $x_1 + x_2 = 1$  är aktivt och att  $x_1 > 0$  och  $x_2 > 0$ .

Dvs.  $g_1(\hat{x}) = 0$ ,  $g_2(\hat{x}) < 0$  och  $g_3(\hat{x}) < 0$ .

KKT2 ger då  $u_2 = 0$  och  $u_3 = 0$ .

$$\text{KKT3 blir då } \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

vilket betyder  $2x_1 - u_1 = 0$  och  $4x_2 - u_1 = 0$ .

Lös ut  $x$  i  $u$ :  $x_1 = u_1/2$  och  $x_2 = u_1/4$ .

Antagandet  $x_1 + x_2 = 1$  ger då  $x_1 + x_2 = u_1/2 + u_1/4 = 1$ ,  
vilket ger  $u_1 = 4/3$ , och  $x_1 = 2/3$  och  $x_2 = 1/3$ .

Detta en KKT-punkt, ty  $u_1 > 0$ ,

## KKT-villkoren: Exempel

Istället för att ange en punkt kan man anta vilka bivillkor som är aktiva.  
(Detta fungerar bara ibland.)

Antag att  $x_1 + x_2 = 1$  är aktivt och att  $x_1 > 0$  och  $x_2 > 0$ .

Dvs.  $g_1(\hat{x}) = 0$ ,  $g_2(\hat{x}) < 0$  och  $g_3(\hat{x}) < 0$ .

KKT2 ger då  $u_2 = 0$  och  $u_3 = 0$ .

$$\text{KKT3 blir då } \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix} + u_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

vilket betyder  $2x_1 - u_1 = 0$  och  $4x_2 - u_1 = 0$ .

Lös ut  $x$  i  $u$ :  $x_1 = u_1/2$  och  $x_2 = u_1/4$ .

Antagandet  $x_1 + x_2 = 1$  ger då  $x_1 + x_2 = u_1/2 + u_1/4 = 1$ ,

vilket ger  $u_1 = 4/3$ , och  $x_1 = 2/3$  och  $x_2 = 1/3$ .

Detta en KKT-punkt, ty  $u_1 > 0$ , och globalt optimum, ty problemet är konvext.

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

(Bivillkoren  $x_1 \leq 1$  och  $x_2 \leq 1$  redundanta. Ta bort.)

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

(Bivillkoren  $x_1 \leq 1$  och  $x_2 \leq 1$  redundanta. Ta bort.)

Konvext.

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

(Bivillkoren  $x_1 \leq 1$  och  $x_2 \leq 1$  redundanta. Ta bort.)

Konvext.

Är det optimalt att bara använda en sort?

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

(Bivillkoren  $x_1 \leq 1$  och  $x_2 \leq 1$  redundanta. Ta bort.)

Konvext.

Är det optimalt att bara använda en sort?

Kolla punkterna  $(1,0)$  och  $(0,1)$ .

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

(Bivillkoren  $x_1 \leq 1$  och  $x_2 \leq 1$  redundanta. Ta bort.)

Konvext.

Är det optimalt att bara använda en sort?

Kolla punkterna  $(1,0)$  och  $(0,1)$ .

Först  $\hat{x} = (1,0)$ :



## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

(Bivillkoren  $x_1 \leq 1$  och  $x_2 \leq 1$  redundanta. Ta bort.)

Konvext.

Är det optimalt att bara använda en sort?

Kolla punkterna  $(1,0)$  och  $(0,1)$ .

Först  $\hat{x} = (1,0)$ :

KKT1: Punkten är tillåten.

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

(Bivillkoren  $x_1 \leq 1$  och  $x_2 \leq 1$  redundanta. Ta bort.)

Konvext.

Är det optimalt att bara använda en sort?

Kolla punkterna  $(1,0)$  och  $(0,1)$ .

Först  $\hat{x} = (1,0)$ :

KKT1: Punkten är tillåten.

KKT2: Bivillkor 1 och 3 aktiva, bivillkor 2 inte.  $\Rightarrow u_2 = 0$ .

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

(Bivillkoren  $x_1 \leq 1$  och  $x_2 \leq 1$  redundanta. Ta bort.)

Konvext.

Är det optimalt att bara använda en sort?

Kolla punkterna  $(1,0)$  och  $(0,1)$ .

Först  $\hat{x} = (1, 0)$ :

KKT1: Punkten är tillåten.

KKT2: Bivillkor 1 och 3 aktiva, bivillkor 2 inte.  $\Rightarrow u_2 = 0$ .

$$\text{KKT3: } \begin{pmatrix} 0 \\ -1 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_3 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

(Bivillkoren  $x_1 \leq 1$  och  $x_2 \leq 1$  redundanta. Ta bort.)

Konvext.

Är det optimalt att bara använda en sort?

Kolla punkterna  $(1,0)$  och  $(0,1)$ .

Först  $\hat{x} = (1,0)$ :

KKT1: Punkten är tillåten.

KKT2: Bivillkor 1 och 3 aktiva, bivillkor 2 inte.  $\Rightarrow u_2 = 0$ .

$$\text{KKT3: } \begin{pmatrix} 0 \\ -1 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_3 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $u_1 = 0$ ,  $u_1 - u_3 = 1$ ,

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

(Bivillkoren  $x_1 \leq 1$  och  $x_2 \leq 1$  redundanta. Ta bort.)

Konvext.

Är det optimalt att bara använda en sort?

Kolla punkterna  $(1,0)$  och  $(0,1)$ .

Först  $\hat{x} = (1, 0)$ :

KKT1: Punkten är tillåten.

KKT2: Bivillkor 1 och 3 aktiva, bivillkor 2 inte.  $\Rightarrow u_2 = 0$ .

$$\text{KKT3: } \begin{pmatrix} 0 \\ -1 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_3 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $u_1 = 0$ ,  $u_1 - u_3 = 1$ , vilket har lösningen  $u_1 = 0$  och  $u_3 = -1$ .

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

(Bivillkoren  $x_1 \leq 1$  och  $x_2 \leq 1$  redundanta. Ta bort.)

Konvext.

Är det optimalt att bara använda en sort?

Kolla punkterna  $(1,0)$  och  $(0,1)$ .

Först  $\hat{x} = (1, 0)$ :

KKT1: Punkten är tillåten.

KKT2: Bivillkor 1 och 3 aktiva, bivillkor 2 inte.  $\Rightarrow u_2 = 0$ .

$$\text{KKT3: } \begin{pmatrix} 0 \\ -1 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_3 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $u_1 = 0$ ,  $u_1 - u_3 = 1$ , vilket har lösningen  $u_1 = 0$  och  $u_3 = -1$ .

$\hat{x}$  är ingen KKT-punkt, ty  $u_3 < 0$ .

## Frikeller-exempel

$$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2 \text{ då } x_1 + x_2 \leq 1, x_1 \geq 0 \text{ och } x_2 \geq 0.$$

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Nu  $\hat{x} = (0, 1)$ :



## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Nu  $\hat{x} = (0, 1)$ :

KKT1: Punkten är tillåten.

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Nu  $\hat{x} = (0, 1)$ :

KKT1: Punkten är tillåten.

KKT2: Bivillkor 1 och 2 aktiva, bivillkor 3 inte.  $\Rightarrow u_3 = 0$ .

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Nu  $\hat{x} = (0, 1)$ :

KKT1: Punkten är tillåten.

KKT2: Bivillkor 1 och 2 aktiva, bivillkor 3 inte.  $\Rightarrow u_3 = 0$ .

$$\text{KKT3: } \begin{pmatrix} -2 \\ 3 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Nu  $\hat{x} = (0, 1)$ :

KKT1: Punkten är tillåten.

KKT2: Bivillkor 1 och 2 aktiva, bivillkor 3 inte.  $\Rightarrow u_3 = 0$ .

$$\text{KKT3: } \begin{pmatrix} -2 \\ 3 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $u_1 - u_2 = 2$ ,  $u_1 = -3$ ,

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Nu  $\hat{x} = (0, 1)$ :

KKT1: Punkten är tillåten.

KKT2: Bivillkor 1 och 2 aktiva, bivillkor 3 inte.  $\Rightarrow u_3 = 0$ .

$$\text{KKT3: } \begin{pmatrix} -2 \\ 3 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $u_1 - u_2 = 2$ ,  $u_1 = -3$ , vilket har lösningen  $u_1 = -3$  och  $u_2 = -5$ .

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Nu  $\hat{x} = (0, 1)$ :

KKT1: Punkten är tillåten.

KKT2: Bivillkor 1 och 2 aktiva, bivillkor 3 inte.  $\Rightarrow u_3 = 0$ .

$$\text{KKT3: } \begin{pmatrix} -2 \\ 3 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $u_1 - u_2 = 2$ ,  $u_1 = -3$ , vilket har lösningen  $u_1 = -3$  och  $u_2 = -5$ .

$\hat{x}$  är ingen KKT-punkt, ty  $u_1 < 0$  och  $u_2 < 0$ .

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Antag istället att  $0 < x_1 < 1$ ,  $0 < x_2 < 1$  och  $x_1 + x_2 = 1$ .



## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Antag istället att  $0 < x_1 < 1$ ,  $0 < x_2 < 1$  och  $x_1 + x_2 = 1$ .

KKT2: Bivillkor 1 aktivt, bivillkor 2 och 3 inte.  $\Rightarrow u_2 = 0$ ,  $u_3 = 0$ .

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Antag istället att  $0 < x_1 < 1$ ,  $0 < x_2 < 1$  och  $x_1 + x_2 = 1$ .

KKT2: Bivillkor 1 aktivt, bivillkor 2 och 3 inte.  $\Rightarrow u_2 = 0$ ,  $u_3 = 0$ .

$$\text{KKT3: } \begin{pmatrix} 2x_1 - 2 \\ 4x_2 - 1 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Antag istället att  $0 < x_1 < 1$ ,  $0 < x_2 < 1$  och  $x_1 + x_2 = 1$ .

KKT2: Bivillkor 1 aktivt, bivillkor 2 och 3 inte.  $\Rightarrow u_2 = 0$ ,  $u_3 = 0$ .

$$\text{KKT3: } \begin{pmatrix} 2x_1 - 2 \\ 4x_2 - 1 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $2x_1 + u_1 = 2$ ,  $4x_2 + u_1 = 1$ ,

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Antag istället att  $0 < x_1 < 1$ ,  $0 < x_2 < 1$  och  $x_1 + x_2 = 1$ .

KKT2: Bivillkor 1 aktivt, bivillkor 2 och 3 inte.  $\Rightarrow u_2 = 0$ ,  $u_3 = 0$ .

$$\text{KKT3: } \begin{pmatrix} 2x_1 - 2 \\ 4x_2 - 1 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $2x_1 + u_1 = 2$ ,  $4x_2 + u_1 = 1$ , vilket ger  $x_1 = (2 - u_1)/2$  och  $x_2 = (1 - u_1)/4$ .

## Frikeller-exempel

min  $f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Antag istället att  $0 < x_1 < 1$ ,  $0 < x_2 < 1$  och  $x_1 + x_2 = 1$ .

KKT2: Bivillkor 1 aktivt, bivillkor 2 och 3 inte.  $\Rightarrow u_2 = 0$ ,  $u_3 = 0$ .

$$\text{KKT3: } \begin{pmatrix} 2x_1 - 2 \\ 4x_2 - 1 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $2x_1 + u_1 = 2$ ,  $4x_2 + u_1 = 1$ , vilket ger  $x_1 = (2 - u_1)/2$  och  $x_2 = (1 - u_1)/4$ .

KKT1:  $x_1 + x_2 = 1$  ger  $(2 - u_1)/2 + (1 - u_1)/4 = 1$ , vilket ger  $u_1 = 1/3$ .

## Frikeller-exempel

min  $f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Antag istället att  $0 < x_1 < 1$ ,  $0 < x_2 < 1$  och  $x_1 + x_2 = 1$ .

KKT2: Bivillkor 1 aktivt, bivillkor 2 och 3 inte.  $\Rightarrow u_2 = 0$ ,  $u_3 = 0$ .

$$\text{KKT3: } \begin{pmatrix} 2x_1 - 2 \\ 4x_2 - 1 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $2x_1 + u_1 = 2$ ,  $4x_2 + u_1 = 1$ , vilket ger  $x_1 = (2 - u_1)/2$  och  $x_2 = (1 - u_1)/4$ .

KKT1:  $x_1 + x_2 = 1$  ger  $(2 - u_1)/2 + (1 - u_1)/4 = 1$ , vilket ger  $u_1 = 1/3$ .

Det ger  $x_1 = 5/6$  och  $x_2 = 1/6$ .

## Frikeller-exempel

min  $f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Antag istället att  $0 < x_1 < 1$ ,  $0 < x_2 < 1$  och  $x_1 + x_2 = 1$ .

KKT2: Bivillkor 1 aktivt, bivillkor 2 och 3 inte.  $\Rightarrow u_2 = 0$ ,  $u_3 = 0$ .

$$\text{KKT3: } \begin{pmatrix} 2x_1 - 2 \\ 4x_2 - 1 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $2x_1 + u_1 = 2$ ,  $4x_2 + u_1 = 1$ , vilket ger  $x_1 = (2 - u_1)/2$  och  $x_2 = (1 - u_1)/4$ .

KKT1:  $x_1 + x_2 = 1$  ger  $(2 - u_1)/2 + (1 - u_1)/4 = 1$ , vilket ger  $u_1 = 1/3$ .

Det ger  $x_1 = 5/6$  och  $x_2 = 1/6$ .

$x_1 \geq 0$  och  $x_2 \geq 0$  så KKT1 är uppfyllt.

## Frikeller-exempel

min  $f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Antag istället att  $0 < x_1 < 1$ ,  $0 < x_2 < 1$  och  $x_1 + x_2 = 1$ .

KKT2: Bivillkor 1 aktivt, bivillkor 2 och 3 inte.  $\Rightarrow u_2 = 0$ ,  $u_3 = 0$ .

$$\text{KKT3: } \begin{pmatrix} 2x_1 - 2 \\ 4x_2 - 1 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $2x_1 + u_1 = 2$ ,  $4x_2 + u_1 = 1$ , vilket ger  $x_1 = (2 - u_1)/2$  och  $x_2 = (1 - u_1)/4$ .

KKT1:  $x_1 + x_2 = 1$  ger  $(2 - u_1)/2 + (1 - u_1)/4 = 1$ , vilket ger  $u_1 = 1/3$ .

Det ger  $x_1 = 5/6$  och  $x_2 = 1/6$ .

$x_1 \geq 0$  och  $x_2 \geq 0$  så KKT1 är uppfyllt.

Detta är en KKT-punkt.



## Frikeller-exempel

min  $f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Antag istället att  $0 < x_1 < 1$ ,  $0 < x_2 < 1$  och  $x_1 + x_2 = 1$ .

KKT2: Bivillkor 1 aktivt, bivillkor 2 och 3 inte.  $\Rightarrow u_2 = 0$ ,  $u_3 = 0$ .

$$\text{KKT3: } \begin{pmatrix} 2x_1 - 2 \\ 4x_2 - 1 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $2x_1 + u_1 = 2$ ,  $4x_2 + u_1 = 1$ , vilket ger  $x_1 = (2 - u_1)/2$  och  $x_2 = (1 - u_1)/4$ .

KKT1:  $x_1 + x_2 = 1$  ger  $(2 - u_1)/2 + (1 - u_1)/4 = 1$ , vilket ger  $u_1 = 1/3$ .

Det ger  $x_1 = 5/6$  och  $x_2 = 1/6$ .

$x_1 \geq 0$  och  $x_2 \geq 0$  så KKT1 är uppfyllt.

Detta är en KKT-punkt.

Optimum, ty problemet konvext.

## Frikeller-exempel

$\min f(x) = x_1^2 + 2x_2^2 - 2x_1 - x_2$  då  $x_1 + x_2 \leq 1$ ,  $x_1 \geq 0$  och  $x_2 \geq 0$ .

Antag istället att  $0 < x_1 < 1$ ,  $0 < x_2 < 1$  och  $x_1 + x_2 = 1$ .

KKT2: Bivillkor 1 aktivt, bivillkor 2 och 3 inte.  $\Rightarrow u_2 = 0$ ,  $u_3 = 0$ .

$$\text{KKT3: } \begin{pmatrix} 2x_1 - 2 \\ 4x_2 - 1 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs.  $2x_1 + u_1 = 2$ ,  $4x_2 + u_1 = 1$ , vilket ger  $x_1 = (2 - u_1)/2$  och  $x_2 = (1 - u_1)/4$ .

KKT1:  $x_1 + x_2 = 1$  ger  $(2 - u_1)/2 + (1 - u_1)/4 = 1$ , vilket ger  $u_1 = 1/3$ .

Det ger  $x_1 = 5/6$  och  $x_2 = 1/6$ .

$x_1 \geq 0$  och  $x_2 \geq 0$  så KKT1 är uppfyllt.

Detta är en KKT-punkt.

Optimum, ty problemet konvext.

Svar: Använd proportionerna  $5/6$  Amarillo och  $1/6$  Citra.

# Olinjär optimering utan bivillkor: Metoder

# Olinjär optimering utan bivillkor: Metoder

$$\min f(x)$$

# Olinjär optimering utan bivillkor: Metoder

$$\min f(x)$$

## Generell sökmetod:

- 1 Finn en startpunkt,  $x^{(k)}$ . Sätt  $k = 1$ .

# Olinjär optimering utan bivillkor: Metoder

$$\min f(x)$$

## Generell sökmetod:

- 1 Finn en startpunkt,  $x^{(k)}$ . Sätt  $k = 1$ .
- 2 Beräkna en **sökriktning**,  $d^{(k)}$ .

# Olinjär optimering utan bivillkor: Metoder

$$\min f(x)$$

## Generell sökmetod:

- 1 Finn en startpunkt,  $x^{(k)}$ . Sätt  $k = 1$ .
- 2 Beräkna en **sökriktning**,  $d^{(k)}$ .
- 3 Om  $\|d^{(k)}\| \leq \varepsilon$  stopp.

# Olinjär optimering utan bivillkor: Metoder

$$\min f(x)$$

## Generell sökmetod:

- 1 Finn en startpunkt,  $x^{(k)}$ . Sätt  $k = 1$ .
- 2 Beräkna en **sökriktning**,  $d^{(k)}$ .
- 3 Om  $\|d^{(k)}\| \leq \varepsilon$  stopp.
- 4 Beräkna en **steglängd**,  $t^{(k)}$ .



# Olinjär optimering utan bivillkor: Metoder

$$\min f(x)$$

## Generell sökmetod:

- 1 Finn en startpunkt,  $x^{(k)}$ . Sätt  $k = 1$ .
- 2 Beräkna en **sökriktning**,  $d^{(k)}$ .
- 3 Om  $\|d^{(k)}\| \leq \varepsilon$  stopp.
- 4 Beräkna en **steglängd**,  $t^{(k)}$ .
- 5 Om  $t^{(k)} \leq \varepsilon$  stopp.

# Olinjär optimering utan bivillkor: Metoder

$$\min f(x)$$

## Generell sökmetod:

- 1 Finn en startpunkt,  $x^{(k)}$ . Sätt  $k = 1$ .
- 2 Beräkna en **sökriktning**,  $d^{(k)}$ .
- 3 Om  $\|d^{(k)}\| \leq \varepsilon$  stopp.
- 4 Beräkna en **steglängd**,  $t^{(k)}$ .
- 5 Om  $t^{(k)} \leq \varepsilon$  stopp.
- 6 Sätt  $x^{(k+1)} = x^{(k)} + t^{(k)}d^{(k)}$ . Sätt  $k = k + 1$  och gå till 2.

# Olinjär optimering utan bivillkor: Metoder

$$\min f(x)$$

## Generell sökmetod:

- 1 Finn en startpunkt,  $x^{(k)}$ . Sätt  $k = 1$ .
- 2 Beräkna en **sökriktning**,  $d^{(k)}$ .
- 3 Om  $\|d^{(k)}\| \leq \varepsilon$  stopp.
- 4 Beräkna en **steglängd**,  $t^{(k)}$ .
- 5 Om  $t^{(k)} \leq \varepsilon$  stopp.
- 6 Sätt  $x^{(k+1)} = x^{(k)} + t^{(k)}d^{(k)}$ . Sätt  $k = k + 1$  och gå till 2.

# Olinjär optimering utan bivillkor: Metoder

$$\min f(x)$$

## Generell sökmetod:

- 1 Finn en startpunkt,  $x^{(k)}$ . Sätt  $k = 1$ .
- 2 Beräkna en **sökriktning**,  $d^{(k)}$ .
- 3 Om  $\|d^{(k)}\| \leq \varepsilon$  stopp.
- 4 Beräkna en **steglängd**,  $t^{(k)}$ .
- 5 Om  $t^{(k)} \leq \varepsilon$  stopp.
- 6 Sätt  $x^{(k+1)} = x^{(k)} + t^{(k)}d^{(k)}$ . Sätt  $k = k + 1$  och gå till 2.

Steglängd: För en given riktning bestäms  $t$  med **linjesökning**

# Olinjär optimering utan bivillkor: Metoder

$$\min f(x)$$

## Generell sökmetod:

- 1 Finn en startpunkt,  $x^{(k)}$ . Sätt  $k = 1$ .
- 2 Beräkna en **sökriktning**,  $d^{(k)}$ .
- 3 Om  $\|d^{(k)}\| \leq \varepsilon$  stopp.
- 4 Beräkna en **steglängd**,  $t^{(k)}$ .
- 5 Om  $t^{(k)} \leq \varepsilon$  stopp.
- 6 Sätt  $x^{(k+1)} = x^{(k)} + t^{(k)}d^{(k)}$ . Sätt  $k = k + 1$  och gå till 2.

Steglängd: För en given riktning bestäms  $t$  med **linjesökning** (en-dimensionell optimering):

# Olinjär optimering utan bivillkor: Metoder

$$\min f(x)$$

## Generell sökmetod:

- 1 Finn en startpunkt,  $x^{(k)}$ . Sätt  $k = 1$ .
- 2 Beräkna en **sökriktning**,  $d^{(k)}$ .
- 3 Om  $\|d^{(k)}\| \leq \varepsilon$  stopp.
- 4 Beräkna en **steglängd**,  $t^{(k)}$ .
- 5 Om  $t^{(k)} \leq \varepsilon$  stopp.
- 6 Sätt  $x^{(k+1)} = x^{(k)} + t^{(k)}d^{(k)}$ . Sätt  $k = k + 1$  och gå till 2.

Steglängd: För en given riktning bestäms  $t$  med **linjesökning** (en-dimensionell optimering):

$$\text{Finn } t^{(k)} \text{ ur } \min_{t \geq 0} f(x^{(k)} + td^{(k)}).$$

# Olinjär optimering utan bivillkor: Metoder

$$\min f(x)$$

## Generell sökmetod:

- 1 Finn en startpunkt,  $x^{(k)}$ . Sätt  $k = 1$ .
- 2 Beräkna en **sökriktning**,  $d^{(k)}$ .
- 3 Om  $\|d^{(k)}\| \leq \varepsilon$  stopp.
- 4 Beräkna en **steglängd**,  $t^{(k)}$ .
- 5 Om  $t^{(k)} \leq \varepsilon$  stopp.
- 6 Sätt  $x^{(k+1)} = x^{(k)} + t^{(k)}d^{(k)}$ . Sätt  $k = k + 1$  och gå till 2.

Steglängd: För en given riktning bestäms  $t$  med **linjesökning** (en-dimensionell optimering):

$$\text{Finn } t^{(k)} \text{ ur } \min_{t \geq 0} f(x^{(k)} + td^{(k)}).$$

(I denna kurs analytiskt.)

## Olinjär optimering utan bivillkor: Linjesökning

Exempel på linjesökning för funktionen  $f(x) = 2x_1^2 + x_2^2 - 3x_1 + x_2$   
i punkten  $\hat{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$  och riktningen  $d = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ .



## Olinjär optimering utan bivillkor: Linjesökning

Exempel på linjesökning för funktionen  $f(x) = 2x_1^2 + x_2^2 - 3x_1 + x_2$  i punkten  $\hat{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$  och riktningen  $d = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ .

Vår nya punkt blir  $x(t) = \hat{x} + td = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + t \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 - t \\ 2 - t \end{pmatrix}$ .

## Olinjär optimering utan bivillkor: Linjesökning

Exempel på linjesökning för funktionen  $f(x) = 2x_1^2 + x_2^2 - 3x_1 + x_2$  i punkten  $\hat{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$  och riktningen  $d = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ .

Vår nya punkt blir  $x(t) = \hat{x} + td = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + t \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1-t \\ 2-t \end{pmatrix}$ .

Sätt in i funktionen:

$$\begin{aligned} f(x(t)) &= 2(1-t)^2 + (2-t)^2 - 3(1-t) + (2-t) = 2(1+t^2-2t) + (4+t^2-4t) - 3 + 3t + 2 - t \\ &= 2 + 2t^2 - 4t + 4 + t^2 - 4t - 3 + 3t + 2 - t = 3t^2 - 6t + 5. \end{aligned}$$

## Olinjär optimering utan bivillkor: Linjesökning

Exempel på linjesökning för funktionen  $f(x) = 2x_1^2 + x_2^2 - 3x_1 + x_2$  i punkten  $\hat{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$  och riktningen  $d = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ .

Vår nya punkt blir  $x(t) = \hat{x} + td = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + t \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1-t \\ 2-t \end{pmatrix}$ .

Sätt in i funktionen:

$$f(x(t)) = 2(1-t)^2 + (2-t)^2 - 3(1-t) + (2-t) = 2(1+t^2-2t) + (4+t^2-4t) - 3 + 3t + 2 - t = 2 + 2t^2 - 4t + 4 + t^2 - 4t - 3 + 3t + 2 - t = 3t^2 - 6t + 5.$$

Derivera:  $f'(x(t)) = 6t - 6$ .

## Olinjär optimering utan bivillkor: Linjesökning

Exempel på linjesökning för funktionen  $f(x) = 2x_1^2 + x_2^2 - 3x_1 + x_2$  i punkten  $\hat{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$  och riktningen  $d = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ .

Vår nya punkt blir  $x(t) = \hat{x} + td = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + t \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1-t \\ 2-t \end{pmatrix}$ .

Sätt in i funktionen:

$$f(x(t)) = 2(1-t)^2 + (2-t)^2 - 3(1-t) + (2-t) = 2(1+t^2-2t) + (4+t^2-4t) - 3 + 3t + 2 - t = 2 + 2t^2 - 4t + 4 + t^2 - 4t - 3 + 3t + 2 - t = 3t^2 - 6t + 5.$$

Derivera:  $f'(x(t)) = 6t - 6$ .

Sätt derivatan till noll (om funktionen är konvex):  $f'(x(t)) = 6t - 6 = 0$  ger  $t = 1$ .

## Olinjär optimering utan bivillkor: Linjesökning

Exempel på linjesökning för funktionen  $f(x) = 2x_1^2 + x_2^2 - 3x_1 + x_2$  i punkten  $\hat{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$  och riktningen  $d = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ .

Vår nya punkt blir  $x(t) = \hat{x} + td = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + t \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1-t \\ 2-t \end{pmatrix}$ .

Sätt in i funktionen:

$$f(x(t)) = 2(1-t)^2 + (2-t)^2 - 3(1-t) + (2-t) = 2(1+t^2-2t) + (4+t^2-4t) - 3 + 3t + 2 - t = 2 + 2t^2 - 4t + 4 + t^2 - 4t - 3 + 3t + 2 - t = 3t^2 - 6t + 5.$$

Derivera:  $f'(x(t)) = 6t - 6$ .

Sätt derivatan till noll (om funktionen är konvex):  $f'(x(t)) = 6t - 6 = 0$  ger  $t = 1$ .

Sätt in i nya punkten:  $x(1) = \begin{pmatrix} 1-1 \\ 2-1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ .

# Grundläggande lokala sökmetoder

Sökriktningen kan beräknas på flera olika sätt.

## Grundläggande lokala sökmetoder

Sökriktningen kan beräknas på flera olika sätt.

Gradienten  $\nabla f(x)$  pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

## Grundläggande lokala sökmetoder

Sökriktningen kan beräknas på flera olika sätt.

Gradienten  $\nabla f(x)$  pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Vill man minimera  $f(x)$  verkar  $-\nabla f(x)$  vara en bra riktning att gå i.



## Grundläggande lokala sökmetoder

Sökriktningen kan beräknas på flera olika sätt.

Gradienten  $\nabla f(x)$  pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Vill man minimera  $f(x)$  verkar  $-\nabla f(x)$  vara en bra riktning att gå i.

**Brantaste lutningsmetoden:** Sätt  $d^{(k)} = -\nabla f(x^{(k)})$ .

# Grundläggande lokala sökmetoder

Sökriktningen kan beräknas på flera olika sätt.

Gradienten  $\nabla f(x)$  pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Vill man minimera  $f(x)$  verkar  $-\nabla f(x)$  vara en bra riktning att gå i.

**Brantaste lutningsmetoden:** Sätt  $d^{(k)} = -\nabla f(x^{(k)})$ .

Ta med andra ordningens information

## Grundläggande lokala sökmetoder

Sökriktningen kan beräknas på flera olika sätt.

Gradienten  $\nabla f(x)$  pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Vill man minimera  $f(x)$  verkar  $-\nabla f(x)$  vara en bra riktning att gå i.

**Brantaste lutningsmetoden:** Sätt  $d^{(k)} = -\nabla f(x^{(k)})$ .

Ta med andra ordningens information (dvs. sikta mot minimum av andra ordningens approximation av  $f(x)$ ):

# Grundläggande lokala sökmetoder

Sökriktningen kan beräknas på flera olika sätt.

Gradienten  $\nabla f(x)$  pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Vill man minimera  $f(x)$  verkar  $-\nabla f(x)$  vara en bra riktning att gå i.

**Brantaste lutningsmetoden:** Sätt  $d^{(k)} = -\nabla f(x^{(k)})$ .

Ta med andra ordningens information (dvs. sikta mot minimum av andra ordningens approximation av  $f(x)$ ):

**Newtons metod:** Sätt  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

# Grundläggande lokala sökmetoder

Sökriktningen kan beräknas på flera olika sätt.

Gradienten  $\nabla f(x)$  pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Vill man minimera  $f(x)$  verkar  $-\nabla f(x)$  vara en bra riktning att gå i.

**Brantaste lutningsmetoden:** Sätt  $d^{(k)} = -\nabla f(x^{(k)})$ .

Ta med andra ordningens information (dvs. sikta mot minimum av andra ordningens approximation av  $f(x)$ ):

**Newtons metod:** Sätt  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Brantaste lutningsmetoden är robust,

# Grundläggande lokala sökmetoder

Sökriktningen kan beräknas på flera olika sätt.

Gradienten  $\nabla f(x)$  pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Vill man minimera  $f(x)$  verkar  $-\nabla f(x)$  vara en bra riktning att gå i.

**Brantaste lutningsmetoden:** Sätt  $d^{(k)} = -\nabla f(x^{(k)})$ .

Ta med andra ordningens information (dvs. sikta mot minimum av andra ordningens approximation av  $f(x)$ ):

**Newtons metod:** Sätt  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Brantaste lutningsmetoden är robust, men kan sicksacka. (lab 3)

# Grundläggande lokala sökmetoder

Sökriktningen kan beräknas på flera olika sätt.

Gradienten  $\nabla f(x)$  pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Vill man minimera  $f(x)$  verkar  $-\nabla f(x)$  vara en bra riktning att gå i.

**Brantaste lutningsmetoden:** Sätt  $d^{(k)} = -\nabla f(x^{(k)})$ .

Ta med andra ordningens information (dvs. sikta mot minimum av andra ordningens approximation av  $f(x)$ ):

**Newtons metod:** Sätt  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Brantaste lutningsmetoden är robust, men kan sicksacka. (lab 3)

Newtons metod använder andra ordningens approximation

# Grundläggande lokala sökmetoder

Sökriktningen kan beräknas på flera olika sätt.

Gradienten  $\nabla f(x)$  pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Vill man minimera  $f(x)$  verkar  $-\nabla f(x)$  vara en bra riktning att gå i.

**Brantaste lutningsmetoden:** Sätt  $d^{(k)} = -\nabla f(x^{(k)})$ .

Ta med andra ordningens information (dvs. sikta mot minimum av andra ordningens approximation av  $f(x)$ ):

**Newtons metod:** Sätt  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Brantaste lutningsmetoden är robust, men kan sicksacka. (lab 3)

Newtons metod använder andra ordningens approximation och är generellt sett starkare



# Grundläggande lokala sökmetoder

Sökriktningen kan beräknas på flera olika sätt.

Gradienten  $\nabla f(x)$  pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Vill man minimera  $f(x)$  verkar  $-\nabla f(x)$  vara en bra riktning att gå i.

**Brantaste lutningsmetoden:** Sätt  $d^{(k)} = -\nabla f(x^{(k)})$ .

Ta med andra ordningens information (dvs. sikta mot minimum av andra ordningens approximation av  $f(x)$ ):

**Newtons metod:** Sätt  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Brantaste lutningsmetoden är robust, men kan sicksacka. (lab 3)

Newtons metod använder andra ordningens approximation och är generellt sett starkare men jobbigare

# Grundläggande lokala sökmetoder

Sökriktningen kan beräknas på flera olika sätt.

Gradienten  $\nabla f(x)$  pekar i den riktning där funktionen  $f(x)$  ökar snabbast.

Vill man minimera  $f(x)$  verkar  $-\nabla f(x)$  vara en bra riktning att gå i.

**Brantaste lutningsmetoden:** Sätt  $d^{(k)} = -\nabla f(x^{(k)})$ .

Ta med andra ordningens information (dvs. sikta mot minimum av andra ordningens approximation av  $f(x)$ ):

**Newtons metod:** Sätt  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Brantaste lutningsmetoden är robust, men kan sicksacka. (lab 3)

Newtons metod använder andra ordningens approximation och är generellt sett starkare men jobbigare och kan falla utan konvexitet.

## Lokala sökmetoder: Exempel

Exempel:  $\min 3x_1^2 + 2x_2^2 - x_2x_2 - 3x_1 - 4x_2$ .

## Lokala sökmetoder: Exempel

Exempel:  $\min 3x_1^2 + 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är

$$\nabla f(x) = \begin{pmatrix} 6x_1 - x_2 - 3 \\ 4x_2 - x_1 - 4 \end{pmatrix}.$$

## Lokala sökmetoder: Exempel

Exempel:  $\min 3x_1^2 + 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är

$$\nabla f(x) = \begin{pmatrix} 6x_1 - x_2 - 3 \\ 4x_2 - x_1 - 4 \end{pmatrix}. \text{ Hessianen är } H(x) = \begin{pmatrix} 6 & -1 \\ -1 & 4 \end{pmatrix}.$$

## Lokala sökmetoder: Exempel

Exempel:  $\min 3x_1^2 + 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är

$$\nabla f(x) = \begin{pmatrix} 6x_1 - x_2 - 3 \\ 4x_2 - x_1 - 4 \end{pmatrix}. \text{ Hessianen är } H(x) = \begin{pmatrix} 6 & -1 \\ -1 & 4 \end{pmatrix}. \text{ Dess}$$

egenvärden är 3.6 och 6.4, båda positiva,

## Lokala sökmetoder: Exempel

Exempel:  $\min 3x_1^2 + 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är

$\nabla f(x) = \begin{pmatrix} 6x_1 - x_2 - 3 \\ 4x_2 - x_1 - 4 \end{pmatrix}$ . Hessianen är  $H(x) = \begin{pmatrix} 6 & -1 \\ -1 & 4 \end{pmatrix}$ . Dess  
egenvärden är 3.6 och 6.4, båda positiva, så funktionen  $f(x)$  är konvex.

## Lokala sökmetoder: Exempel

Exempel:  $\min 3x_1^2 + 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är

$\nabla f(x) = \begin{pmatrix} 6x_1 - x_2 - 3 \\ 4x_2 - x_1 - 4 \end{pmatrix}$ . Hessianen är  $H(x) = \begin{pmatrix} 6 & -1 \\ -1 & 4 \end{pmatrix}$ . Dess egenvärden är 3.6 och 6.4, båda positiva, så funktionen  $f(x)$  är konvex.

I punkten  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 1$  har vi  $\nabla f(x) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ ,



## Lokala sökmetoder: Exempel

Exempel:  $\min 3x_1^2 + 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är

$\nabla f(x) = \begin{pmatrix} 6x_1 - x_2 - 3 \\ 4x_2 - x_1 - 4 \end{pmatrix}$ . Hessianen är  $H(x) = \begin{pmatrix} 6 & -1 \\ -1 & 4 \end{pmatrix}$ . Dess egenvärden är 3.6 och 6.4, båda positiva, så funktionen  $f(x)$  är konvex.

I punkten  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 1$  har vi  $\nabla f(x) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ , så brantaste

lutningsriktningen är  $d_{br} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ .

## Lokala sökmetoder: Exempel

Exempel:  $\min 3x_1^2 + 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är

$\nabla f(x) = \begin{pmatrix} 6x_1 - x_2 - 3 \\ 4x_2 - x_1 - 4 \end{pmatrix}$ . Hessianen är  $H(x) = \begin{pmatrix} 6 & -1 \\ -1 & 4 \end{pmatrix}$ . Dess egenvärden är 3.6 och 6.4, båda positiva, så funktionen  $f(x)$  är konvex.

I punkten  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 1$  har vi  $\nabla f(x) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ , så brantaste

lutningsriktningen är  $d_{br} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ .

Vi har  $H^{-1} = \frac{1}{23} \begin{pmatrix} 4 & 1 \\ 1 & 6 \end{pmatrix}$ ,

## Lokala sökmetoder: Exempel

Exempel:  $\min 3x_1^2 + 2x_2^2 - x_2x_2 - 3x_1 - 4x_2$ . Gradienten är  $\nabla f(x) = \begin{pmatrix} 6x_1 - x_2 - 3 \\ 4x_2 - x_1 - 4 \end{pmatrix}$ . Hessianen är  $H(x) = \begin{pmatrix} 6 & -1 \\ -1 & 4 \end{pmatrix}$ . Dess egenvärden är 3.6 och 6.4, båda positiva, så funktionen  $f(x)$  är konvex.

I punkten  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 1$  har vi  $\nabla f(x) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ , så brantaste lutningsriktningen är  $d_{br} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ .

Vi har  $H^{-1} = \frac{1}{23} \begin{pmatrix} 4 & 1 \\ 1 & 6 \end{pmatrix}$ , så Newtonriktningen blir

$$d_N = -H^{-1}\nabla f(\hat{x}) = -\frac{1}{23} \begin{pmatrix} 4 & 1 \\ 1 & 6 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} =$$

## Lokala sökmetoder: Exempel

Exempel:  $\min 3x_1^2 + 2x_2^2 - x_2x_2 - 3x_1 - 4x_2$ . Gradienten är  $\nabla f(x) = \begin{pmatrix} 6x_1 - x_2 - 3 \\ 4x_2 - x_1 - 4 \end{pmatrix}$ . Hessianen är  $H(x) = \begin{pmatrix} 6 & -1 \\ -1 & 4 \end{pmatrix}$ . Dess egenvärden är 3.6 och 6.4, båda positiva, så funktionen  $f(x)$  är konvex.

I punkten  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 1$  har vi  $\nabla f(x) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ , så brantaste lutningsriktningen är  $d_{br} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ .

Vi har  $H^{-1} = \frac{1}{23} \begin{pmatrix} 4 & 1 \\ 1 & 6 \end{pmatrix}$ , så Newtonriktningen blir

$$d_N = -H^{-1}\nabla f(\hat{x}) = -\frac{1}{23} \begin{pmatrix} 4 & 1 \\ 1 & 6 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \frac{1}{23} \begin{pmatrix} -7 \\ 4 \end{pmatrix}$$

## Lokala sökmetoder: Exempel

Exempel:  $\min 3x_1^2 + 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är  $\nabla f(x) = \begin{pmatrix} 6x_1 - x_2 - 3 \\ 4x_2 - x_1 - 4 \end{pmatrix}$ . Hessianen är  $H(x) = \begin{pmatrix} 6 & -1 \\ -1 & 4 \end{pmatrix}$ . Dess egenvärden är 3.6 och 6.4, båda positiva, så funktionen  $f(x)$  är konvex.

I punkten  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 1$  har vi  $\nabla f(x) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ , så brantaste lutningsriktningen är  $d_{br} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ .

Vi har  $H^{-1} = \frac{1}{23} \begin{pmatrix} 4 & 1 \\ 1 & 6 \end{pmatrix}$ , så Newtonriktningen blir

$$d_N = -H^{-1}\nabla f(\hat{x}) = -\frac{1}{23} \begin{pmatrix} 4 & 1 \\ 1 & 6 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \frac{1}{23} \begin{pmatrix} -7 \\ 4 \end{pmatrix}$$

Newtonriktningen pekar precis mot optimum (eftersom funktionen är konvex och kvadratisk).

## Lokala sökmetoder: Exempel

Exempel:  $\min 3x_1^2 + 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är  $\nabla f(x) = \begin{pmatrix} 6x_1 - x_2 - 3 \\ 4x_2 - x_1 - 4 \end{pmatrix}$ . Hessianen är  $H(x) = \begin{pmatrix} 6 & -1 \\ -1 & 4 \end{pmatrix}$ . Dess egenvärden är 3.6 och 6.4, båda positiva, så funktionen  $f(x)$  är konvex.

I punkten  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 1$  har vi  $\nabla f(x) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ , så brantaste lutningsriktningen är  $d_{br} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ .

Vi har  $H^{-1} = \frac{1}{23} \begin{pmatrix} 4 & 1 \\ 1 & 6 \end{pmatrix}$ , så Newtonriktningen blir

$$d_N = -H^{-1}\nabla f(\hat{x}) = -\frac{1}{23} \begin{pmatrix} 4 & 1 \\ 1 & 6 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \frac{1}{23} \begin{pmatrix} -7 \\ 4 \end{pmatrix}$$

Newtonriktningen pekar precis mot optimum (eftersom funktionen är konvex och kvadratisk). Optimal steglängd blir då 1.

## Lokala sökmetoder: Exempel

Exempel:  $\min -3x_1^2 - 2x_2^2 - x_2x_2 - 3x_1 - 4x_2$ .

## Lokala sökmetoder: Exempel

Exempel:  $\min -3x_1^2 - 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är

$$\nabla f(x) = \begin{pmatrix} -6x_1 - x_2 - 3 \\ -4x_2 - x_1 - 4 \end{pmatrix}.$$



## Lokala sökmetoder: Exempel

Exempel:  $\min -3x_1^2 - 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är

$$\nabla f(x) = \begin{pmatrix} -6x_1 - x_2 - 3 \\ -4x_2 - x_1 - 4 \end{pmatrix}. \text{ Hessianen är } H(x) = \begin{pmatrix} -6 & -1 \\ -1 & -4 \end{pmatrix}.$$

## Lokala sökmetoder: Exempel

Exempel:  $\min -3x_1^2 - 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är

$$\nabla f(x) = \begin{pmatrix} -6x_1 - x_2 - 3 \\ -4x_2 - x_1 - 4 \end{pmatrix}. \text{ Hessianen är } H(x) = \begin{pmatrix} -6 & -1 \\ -1 & -4 \end{pmatrix}. \text{ Dess}$$

egenvärden är ung  $-6.4$  och  $-3.6$ , så funktionen  $f(x)$  är inte konvex.

## Lokala sökmetoder: Exempel

Exempel:  $\min -3x_1^2 - 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är

$$\nabla f(x) = \begin{pmatrix} -6x_1 - x_2 - 3 \\ -4x_2 - x_1 - 4 \end{pmatrix}. \text{ Hessianen är } H(x) = \begin{pmatrix} -6 & -1 \\ -1 & -4 \end{pmatrix}. \text{ Dess}$$

egenvärden är ung  $-6.4$  och  $-3.6$ , så funktionen  $f(x)$  är inte konvex.

$$\text{I punkten } \hat{x}_1 = 1 \text{ och } \hat{x}_2 = 1 \text{ har vi } \nabla f(x) = \begin{pmatrix} -10 \\ -9 \end{pmatrix},$$

## Lokala sökmetoder: Exempel

Exempel:  $\min -3x_1^2 - 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är  $\nabla f(x) = \begin{pmatrix} -6x_1 - x_2 - 3 \\ -4x_2 - x_1 - 4 \end{pmatrix}$ . Hessianen är  $H(x) = \begin{pmatrix} -6 & -1 \\ -1 & -4 \end{pmatrix}$ . Dess egenvärden är ung  $-6.4$  och  $-3.6$ , så funktionen  $f(x)$  är inte konvex.

I punkten  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 1$  har vi  $\nabla f(x) = \begin{pmatrix} -10 \\ -9 \end{pmatrix}$ , så brantaste lutningsriktningen är  $d_{br} = \begin{pmatrix} 10 \\ 9 \end{pmatrix}$ .

## Lokala sökmetoder: Exempel

Exempel:  $\min -3x_1^2 - 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är  $\nabla f(x) = \begin{pmatrix} -6x_1 - x_2 - 3 \\ -4x_2 - x_1 - 4 \end{pmatrix}$ . Hessianen är  $H(x) = \begin{pmatrix} -6 & -1 \\ -1 & -4 \end{pmatrix}$ . Dess egenvärden är ung  $-6.4$  och  $-3.6$ , så funktionen  $f(x)$  är inte konvex.

I punkten  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 1$  har vi  $\nabla f(x) = \begin{pmatrix} -10 \\ -9 \end{pmatrix}$ , så brantaste

lutningsriktningen är  $d_{br} = \begin{pmatrix} 10 \\ 9 \end{pmatrix}$ .

Vi har  $H^{-1} = \frac{1}{23} \begin{pmatrix} -4 & 1 \\ 1 & -6 \end{pmatrix}$ ,

## Lokala sökmetoder: Exempel

Exempel:  $\min -3x_1^2 - 2x_2^2 - x_2x_2 - 3x_1 - 4x_2$ . Gradienten är  $\nabla f(x) = \begin{pmatrix} -6x_1 - x_2 - 3 \\ -4x_2 - x_1 - 4 \end{pmatrix}$ . Hessianen är  $H(x) = \begin{pmatrix} -6 & -1 \\ -1 & -4 \end{pmatrix}$ . Dess egenvärden är ung  $-6.4$  och  $-3.6$ , så funktionen  $f(x)$  är inte konvex.

I punkten  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 1$  har vi  $\nabla f(x) = \begin{pmatrix} -10 \\ -9 \end{pmatrix}$ , så brantaste lutningsriktningen är  $d_{br} = \begin{pmatrix} 10 \\ 9 \end{pmatrix}$ .

Vi har  $H^{-1} = \frac{1}{23} \begin{pmatrix} -4 & 1 \\ 1 & -6 \end{pmatrix}$ , så Newtonriktningen blir

$$d_N = -H^{-1}\nabla f(\hat{x}) = -\frac{1}{23} \begin{pmatrix} -4 & 1 \\ 1 & -6 \end{pmatrix} \begin{pmatrix} -10 \\ -9 \end{pmatrix} =$$

## Lokala sökmetoder: Exempel

Exempel:  $\min -3x_1^2 - 2x_2^2 - x_2x_2 - 3x_1 - 4x_2$ . Gradienten är  $\nabla f(x) = \begin{pmatrix} -6x_1 - x_2 - 3 \\ -4x_2 - x_1 - 4 \end{pmatrix}$ . Hessianen är  $H(x) = \begin{pmatrix} -6 & -1 \\ -1 & -4 \end{pmatrix}$ . Dess egenvärden är ung  $-6.4$  och  $-3.6$ , så funktionen  $f(x)$  är inte konvex.

I punkten  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 1$  har vi  $\nabla f(x) = \begin{pmatrix} -10 \\ -9 \end{pmatrix}$ , så brantaste lutningsriktningen är  $d_{br} = \begin{pmatrix} 10 \\ 9 \end{pmatrix}$ .

Vi har  $H^{-1} = \frac{1}{23} \begin{pmatrix} -4 & 1 \\ 1 & -6 \end{pmatrix}$ , så Newtonriktningen blir

$$d_N = -H^{-1}\nabla f(\hat{x}) = -\frac{1}{23} \begin{pmatrix} -4 & 1 \\ 1 & -6 \end{pmatrix} \begin{pmatrix} -10 \\ -9 \end{pmatrix} = \frac{1}{23} \begin{pmatrix} -31 \\ -44 \end{pmatrix}$$

## Lokala sökmetoder: Exempel

Exempel:  $\min -3x_1^2 - 2x_2^2 - x_2x_1 - 3x_1 - 4x_2$ . Gradienten är  $\nabla f(x) = \begin{pmatrix} -6x_1 - x_2 - 3 \\ -4x_2 - x_1 - 4 \end{pmatrix}$ . Hessianen är  $H(x) = \begin{pmatrix} -6 & -1 \\ -1 & -4 \end{pmatrix}$ . Dess egenvärden är ung  $-6.4$  och  $-3.6$ , så funktionen  $f(x)$  är inte konvex.

I punkten  $\hat{x}_1 = 1$  och  $\hat{x}_2 = 1$  har vi  $\nabla f(x) = \begin{pmatrix} -10 \\ -9 \end{pmatrix}$ , så brantaste lutningsriktningen är  $d_{br} = \begin{pmatrix} 10 \\ 9 \end{pmatrix}$ .

Vi har  $H^{-1} = \frac{1}{23} \begin{pmatrix} -4 & 1 \\ 1 & -6 \end{pmatrix}$ , så Newtonriktningen blir

$$d_N = -H^{-1}\nabla f(\hat{x}) = -\frac{1}{23} \begin{pmatrix} -4 & 1 \\ 1 & -6 \end{pmatrix} \begin{pmatrix} -10 \\ -9 \end{pmatrix} = \frac{1}{23} \begin{pmatrix} -31 \\ -44 \end{pmatrix}$$

Newtonriktningen är inte en avtaganderiktning, utan en ökanderiktning. Linjesökning skulle ge steg noll.



# Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Newton's metod fungerar dåligt för ickekonvexa funktioner.

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Newton's metod fungerar dåligt för ickekonvexa funktioner.

Kan t.o.m. ge ökningsriktning (istället för avtaganderiktning).

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Newtons metod fungerar dåligt för ickekonvexa funktioner.

Kan t.o.m. ge ökningsriktning (istället för avtaganderiktning).

**Konvexifiering:**

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Newtons metod fungerar dåligt för ickekonvexa funktioner.

Kan t.o.m. ge ökningsriktning (istället för avtaganderiktning).

**Konvexifiering:**

**Levenbergs modifiering** av Newtons smetod: Ersätt  $H$  med  $H + \nu I$ .

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Newtons metod fungerar dåligt för ickekonvexa funktioner.

Kan t.o.m. ge ökningsriktning (istället för avtaganderiktning).

**Konvexifiering:**

**Levenbergs modifiering** av Newtons smetod: Ersätt  $H$  med  $H + \nu I$ .

Som att addera konvexa termen  $\nu(x_1^2 + x_2^2 + \dots + x_n^2)$ .

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Newtons metod fungerar dåligt för ickekonvexa funktioner.

Kan t.o.m. ge ökningsriktning (istället för avtaganderiktning).

**Konvexifiering:**

**Levenbergs modifiering** av Newtons smetod: Ersätt  $H$  med  $H + \nu I$ .

Som att addera konvexa termen  $\nu(x_1^2 + x_2^2 + \dots + x_n^2)$ .

$\nu = 0$  ger Newtons metod.



## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Newtons metod fungerar dåligt för ickekonvexa funktioner.

Kan t.o.m. ge ökningsriktning (istället för avtaganderiktning).

**Konvexifiering:**

**Levenbergs modifiering** av Newtons smetod: Ersätt  $H$  med  $H + \nu I$ .

Som att addera konvexa termen  $\nu(x_1^2 + x_2^2 + \dots + x_n^2)$ .

$\nu = 0$  ger Newtons metod.

Ett mycket stort  $\nu$  ger i princip brantaste lutningsmetoden.

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Newtons metod fungerar dåligt för ickekonvexa funktioner.

Kan t.o.m. ge ökningsriktning (istället för avtaganderiktning).

**Konvexifiering:**

**Levenbergs modifiering** av Newtons smetod: Ersätt  $H$  med  $H + \nu I$ .

Som att addera konvexa termen  $\nu(x_1^2 + x_2^2 + \dots + x_n^2)$ .

$\nu = 0$  ger Newtons metod.

Ett mycket stort  $\nu$  ger i princip brantaste lutningsmetoden.

Större  $\nu$  ger “mer positiva” egenvärden, dvs. en “mer konvex” funktion.

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Newtons metod fungerar dåligt för ickekonvexa funktioner.

Kan t.o.m. ge ökningsriktning (istället för avtaganderiktning).

### Konvexifiering:

**Levenbergs modifiering** av Newtons smetod: Ersätt  $H$  med  $H + \nu I$ .

Som att addera konvexa termen  $\nu(x_1^2 + x_2^2 + \dots + x_n^2)$ .

$\nu = 0$  ger Newtons metod.

Ett mycket stort  $\nu$  ger i princip brantaste lutningsmetoden.

Större  $\nu$  ger "mer positiva" egenvärden, dvs. en "mer konvex" funktion.

Tillräckligt stort  $\nu$  ger konvex funktion.

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Newtons metod fungerar dåligt för ickekonvexa funktioner.

Kan t.o.m. ge ökningsriktning (istället för avtaganderiktning).

### Konvexifiering:

**Levenbergs modifiering** av Newtons smetod: Ersätt  $H$  med  $H + \nu I$ .

Som att addera konvexa termen  $\nu(x_1^2 + x_2^2 + \dots + x_n^2)$ .

$\nu = 0$  ger Newtons metod.

Ett mycket stort  $\nu$  ger i princip brantaste lutningsmetoden.

Större  $\nu$  ger "mer positiva" egenvärden, dvs. en "mer konvex" funktion.

Tillräckligt stort  $\nu$  ger konvex funktion.

Jämför med mest negativa egenvärdet av  $H$ .

# Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Det är jobbigt att beräkna och invertera  $H$  i varje iteration i Newtons metod.

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Det är jobbigt att beräkna och invertera  $H$  i varje iteration i Newtons metod.

**Förenkling av Newtons metod:**



## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Det är jobbigt att beräkna och invertera  $H$  i varje iteration i Newtons metod.

**Förenkling av Newtons metod:**

**Kvasi-Newtonmetoder:** Ersätt  $H^{-1}$  med en enklare beräknad approximation.

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Det är jobbigt att beräkna och invertera  $H$  i varje iteration i Newtons metod.

**Förenkling av Newtons metod:**

**Kvasi-Newtonmetoder:** Ersätt  $H^{-1}$  med en enklare beräknad approximation.

Approximationen uppdateras med enbart gradienter.

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Det är jobbigt att beräkna och invertera  $H$  i varje iteration i Newtons metod.

**Förenkling av Newtons metod:**

**Kvasi-Newtonmetoder:** Ersätt  $H^{-1}$  med en enklare beräknad approximation.

Approximationen uppdateras med enbart gradienter.

Flera varianter finns: DFP (Davidon, Fletcher, Powell), Broyden (Broyden), BFGS (Broyden, Fletcher, Goldfarb, Shanno) och SR1 (symmetric rank 1).

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Det är jobbigt att beräkna och invertera  $H$  i varje iteration i Newtons metod.

**Förenkling av Newtons metod:**

**Kvasi-Newtonmetoder:** Ersätt  $H^{-1}$  med en enklare beräknad approximation.

Approximationen uppdateras med enbart gradienter.

Flera varianter finns: DFP (Davidon, Fletcher, Powell), Broyden (Broyden), BFGS (Broyden, Fletcher, Goldfarb, Shanno) och SR1 (symmetric rank 1).

Krängliga formler för effektiv direkt uppdatering av inversapproximationen.

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Newton's metod:**  $d^{(k)} = -H(x^{(k)})^{-1}\nabla f(x^{(k)})$ .

Det är jobbigt att beräkna och invertera  $H$  i varje iteration i Newtons metod.

**Förenkling av Newtons metod:**

**Kvasi-Newtonmetoder:** Ersätt  $H^{-1}$  med en enklare beräknad approximation.

Approximationen uppdateras med enbart gradienter.

Flera varianter finns: DFP (Davidon, Fletcher, Powell), Broyden (Broyden), BFGS (Broyden, Fletcher, Goldfarb, Shanno) och SR1 (symmetric rank 1).

Krängliga formler för effektiv direkt uppdatering av inversapproximationen.

(Finns i Nileopt.)

# Kvasi-Newtonmetoder

Formler enligt Wikipedia:

Method	$B_{k+1} =$	$H_{k+1} = B_{k+1}^{-1} =$
DFP	$\left(I - \frac{y_k \Delta x_k^T}{y_k^T \Delta x_k}\right) B_k \left(I - \frac{\Delta x_k y_k^T}{y_k^T \Delta x_k}\right) + \frac{y_k y_k^T}{y_k^T \Delta x_k}$	$H_k + \frac{\Delta x_k \Delta x_k^T}{y_k^T \Delta x_k} - \frac{H_k y_k y_k^T H_k^T}{y_k^T H_k y_k}$
BFGS	$B_k + \frac{y_k y_k^T}{y_k^T \Delta x_k} - \frac{B_k \Delta x_k (B_k \Delta x_k)^T}{\Delta x_k^T B_k \Delta x_k}$	$\left(I - \frac{y_k \Delta x_k^T}{y_k^T \Delta x_k}\right)^T H_k \left(I - \frac{y_k \Delta x_k^T}{y_k^T \Delta x_k}\right) + \frac{\Delta x_k \Delta x_k^T}{y_k^T \Delta x_k}$
Broyden	$B_k + \frac{y_k - B_k \Delta x_k}{\Delta x_k^T \Delta x_k} \Delta x_k^T$	$H_k + \frac{(\Delta x_k - H_k y_k) \Delta x_k^T H_k}{\Delta x_k^T H_k y_k}$
Broyden family	$(1 - \varphi_k) B_{k+1}^{BFGS} + \varphi_k B_{k+1}^{DFP}, \quad \varphi \in [0, 1]$	
SR1	$B_k + \frac{(y_k - B_k \Delta x_k)(y_k - B_k \Delta x_k)^T}{(y_k - B_k \Delta x_k)^T \Delta x_k}$	$H_k + \frac{(\Delta x_k - H_k y_k)(\Delta x_k - H_k y_k)^T}{(\Delta x_k - H_k y_k)^T y_k}$

# Kvasi-Newtonmetoder

Formler enligt Wikipedia:

Method	$B_{k+1} =$	$H_{k+1} = B_{k+1}^{-1} =$
DFP	$\left(I - \frac{y_k \Delta x_k^T}{y_k^T \Delta x_k}\right) B_k \left(I - \frac{\Delta x_k y_k^T}{y_k^T \Delta x_k}\right) + \frac{y_k y_k^T}{y_k^T \Delta x_k}$	$H_k + \frac{\Delta x_k \Delta x_k^T}{y_k^T \Delta x_k} - \frac{H_k y_k y_k^T H_k^T}{y_k^T H_k y_k}$
BFGS	$B_k + \frac{y_k y_k^T}{y_k^T \Delta x_k} - \frac{B_k \Delta x_k (B_k \Delta x_k)^T}{\Delta x_k^T B_k \Delta x_k}$	$\left(I - \frac{y_k \Delta x_k^T}{y_k^T \Delta x_k}\right)^T H_k \left(I - \frac{y_k \Delta x_k^T}{y_k^T \Delta x_k}\right) + \frac{\Delta x_k \Delta x_k^T}{y_k^T \Delta x_k}$
Broyden	$B_k + \frac{y_k - B_k \Delta x_k}{\Delta x_k^T \Delta x_k} \Delta x_k^T$	$H_k + \frac{(\Delta x_k - H_k y_k) \Delta x_k^T H_k}{\Delta x_k^T H_k y_k}$
Broyden family	$(1 - \varphi_k) B_{k+1}^{BFGS} + \varphi_k B_{k+1}^{DFP}, \quad \varphi \in [0, 1]$	
SR1	$B_k + \frac{(y_k - B_k \Delta x_k)(y_k - B_k \Delta x_k)^T}{(y_k - B_k \Delta x_k)^T \Delta x_k}$	$H_k + \frac{(\Delta x_k - H_k y_k)(\Delta x_k - H_k y_k)^T}{(\Delta x_k - H_k y_k)^T y_k}$

Vi ska *inte* gå in på detaljer här.

# Kvasi-Newtonmetoder

Formler enligt Wikipedia:

Method	$B_{k+1} =$	$H_{k+1} = B_{k+1}^{-1} =$
DFP	$\left(I - \frac{y_k \Delta x_k^T}{y_k^T \Delta x_k}\right) B_k \left(I - \frac{\Delta x_k y_k^T}{y_k^T \Delta x_k}\right) + \frac{y_k y_k^T}{y_k^T \Delta x_k}$	$H_k + \frac{\Delta x_k \Delta x_k^T}{y_k^T \Delta x_k} - \frac{H_k y_k y_k^T H_k^T}{y_k^T H_k y_k}$
BFGS	$B_k + \frac{y_k y_k^T}{y_k^T \Delta x_k} - \frac{B_k \Delta x_k (B_k \Delta x_k)^T}{\Delta x_k^T B_k \Delta x_k}$	$\left(I - \frac{y_k \Delta x_k^T}{y_k^T \Delta x_k}\right)^T H_k \left(I - \frac{y_k \Delta x_k^T}{y_k^T \Delta x_k}\right) + \frac{\Delta x_k \Delta x_k^T}{y_k^T \Delta x_k}$
Broyden	$B_k + \frac{y_k - B_k \Delta x_k}{\Delta x_k^T \Delta x_k} \Delta x_k^T$	$H_k + \frac{(\Delta x_k - H_k y_k) \Delta x_k^T H_k}{\Delta x_k^T H_k y_k}$
Broyden family	$(1 - \varphi_k) B_{k+1}^{BFGS} + \varphi_k B_{k+1}^{DFP}, \quad \varphi \in [0, 1]$	
SR1	$B_k + \frac{(y_k - B_k \Delta x_k)(y_k - B_k \Delta x_k)^T}{(y_k - B_k \Delta x_k)^T \Delta x_k}$	$H_k + \frac{(\Delta x_k - H_k y_k)(\Delta x_k - H_k y_k)^T}{(\Delta x_k - H_k y_k)^T y_k}$

Vi ska *inte* gå in på detaljer här.

Men de finns i Nileopt. I lab 3 kan ni jämföra dem.



# Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Brantaste lutningsmetoden:**  $d^{(k)} = -\nabla f(x^{(k)})$ .

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Brantaste lutningsmetoden:**  $d^{(k)} = -\nabla f(x^{(k)})$ .

Brantaste lutningsmetoden kan bli långsam p.g.a. sicksackning.

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Brantaste lutningsmetoden:**  $d^{(k)} = -\nabla f(x^{(k)})$ .

Brantaste lutningsmetoden kan bli långsam p.g.a. sicksackning.

Problemet är att närliggande sökriktningar är ortogonala.

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Brantaste lutningsmetoden:**  $d^{(k)} = -\nabla f(x^{(k)})$ .

Brantaste lutningsmetoden kan bli långsam p.g.a. sicksackning.  
Problemet är att närliggande sökriktningar är ortogonala.

**Förbättring av brantaste lutningsmetoden:**

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Brantaste lutningsmetoden:**  $d^{(k)} = -\nabla f(x^{(k)})$ .

Brantaste lutningsmetoden kan bli långsam p.g.a. sicksackning.  
Problemet är att närliggande sökriktningar är ortogonala.

**Förbättring av brantaste lutningsmetoden:**

**Konjugerade gradientmetoder:** Modifiera riktningen för att undvika sicksackning, genom att addera in tidigare sökriktningar (olika mycket).

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Brantaste lutningsmetoden:**  $d^{(k)} = -\nabla f(x^{(k)})$ .

Brantaste lutningsmetoden kan bli långsam p.g.a. sicksackning.

Problemet är att närliggande sökriktningar är ortogonala.

**Förbättring av brantaste lutningsmetoden:**

**Konjugerade gradientmetoder:** Modifiera riktningen för att undvika sicksackning, genom att addera in tidigare sökriktningar (olika mycket).

Sätt  $d^{(k)} = -\nabla f(x^{(k)}) + \beta_k d^{(k-1)}$ .

## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Brantaste lutningsmetoden:**  $d^{(k)} = -\nabla f(x^{(k)})$ .

Brantaste lutningsmetoden kan bli långsam p.g.a. sicksackning.  
Problemet är att närliggande sökriktningar är ortogonala.

**Förbättring av brantaste lutningsmetoden:**

**Konjugerade gradientmetoder:** Modifiera riktningen för att undvika sicksackning, genom att addera in tidigare sökriktningar (olika mycket).

Sätt  $d^{(k)} = -\nabla f(x^{(k)}) + \beta_k d^{(k-1)}$ .

Flera varianter (krångliga formler) för att beräkna  $\beta$  finns: FR (Fletcher-Reeves), HS (Hestenes-Stiefel), PR (Polak-Ribiere).



## Modifierade lokala sökmetoder

Försök åtgärda nackdelar:

**Brantaste lutningsmetoden:**  $d^{(k)} = -\nabla f(x^{(k)})$ .

Brantaste lutningsmetoden kan bli långsam p.g.a. sicksackning.  
Problemet är att närliggande sökriktningar är ortogonala.

**Förbättring av brantaste lutningsmetoden:**

**Konjugerade gradientmetoder:** Modifiera riktningen för att undvika sicksackning, genom att addera in tidigare sökriktningar (olika mycket).

Sätt  $d^{(k)} = -\nabla f(x^{(k)}) + \beta_k d^{(k-1)}$ .

Flera varianter (krångliga formler) för att beräkna  $\beta$  finns: FR (Fletcher-Reeves), HS (Hestenes-Stiefel), PR (Polak-Ribiere).

(Finns i Nileopt.)

## Konjugerade gradientmetoder

**Konjugerade gradientmetoder:** Sätt  $d^{(k)} = -\nabla f(x^{(k)}) + \beta_k d^{(k-1)}$ .

# Konjugerade gradientmetoder

**Konjugerade gradientmetoder:** Sätt  $d^{(k)} = -\nabla f(x^{(k)}) + \beta_k d^{(k-1)}$ .

Formler för att beräkna  $\beta$ :

# Konjugerade gradientmetoder

**Konjugerade gradientmetoder:** Sätt  $d^{(k)} = -\nabla f(x^{(k)}) + \beta_k d^{(k-1)}$ .

Formler för att beräkna  $\beta$ :

Fletcher-Reeves: 
$$\beta_k = \frac{\|\nabla f(x^{(k)})\|^2}{\|\nabla f(x^{(k-1)})\|^2}$$

# Konjugerade gradientmetoder

**Konjugerade gradientmetoder:** Sätt  $d^{(k)} = -\nabla f(x^{(k)}) + \beta_k d^{(k-1)}$ .

Formler för att beräkna  $\beta$ :

Fletcher-Reeves: 
$$\beta_k = \frac{\|\nabla f(x^{(k)})\|^2}{\|\nabla f(x^{(k-1)})\|^2}$$

Polak-Ribiere: 
$$\beta_k = \frac{\nabla f(x^{(k)})^T y^{(k)}}{\|\nabla f(x^{(k-1)})\|^2}$$

# Konjugerade gradientmetoder

**Konjugerade gradientmetoder:** Sätt  $d^{(k)} = -\nabla f(x^{(k)}) + \beta_k d^{(k-1)}$ .

Formler för att beräkna  $\beta$ :

$$\text{Fletcher-Reeves: } \beta_k = \frac{\|\nabla f(x^{(k)})\|^2}{\|\nabla f(x^{(k-1)})\|^2}$$

$$\text{Polak-Ribiere: } \beta_k = \frac{\nabla f(x^{(k)})^T y^{(k)}}{\|\nabla f(x^{(k-1)})\|^2}$$

$$\text{Hestenes-Stiefel: } \beta_k = \frac{\nabla f(x^{(k)})^T y^{(k)}}{d^{(k-1)T} y^{(k)}}$$

# Konjugerade gradientmetoder

**Konjugerade gradientmetoder:** Sätt  $d^{(k)} = -\nabla f(x^{(k)}) + \beta_k d^{(k-1)}$ .

Formler för att beräkna  $\beta$ :

Fletcher-Reeves: 
$$\beta_k = \frac{\|\nabla f(x^{(k)})\|^2}{\|\nabla f(x^{(k-1)})\|^2}$$

Polak-Ribiere: 
$$\beta_k = \frac{\nabla f(x^{(k)})^T y^{(k)}}{\|\nabla f(x^{(k-1)})\|^2}$$

Hestenes-Stiefel: 
$$\beta_k = \frac{\nabla f(x^{(k)})^T y^{(k)}}{d^{(k-1)T} y^{(k)}}$$

där  $y^{(k)} = \nabla f(x^{(k)}) - \nabla f(x^{(k-1)})$ .

# Konjugerade gradientmetoder

**Konjugerade gradientmetoder:** Sätt  $d^{(k)} = -\nabla f(x^{(k)}) + \beta_k d^{(k-1)}$ .

Formler för att beräkna  $\beta$ :

Fletcher-Reeves: 
$$\beta_k = \frac{\|\nabla f(x^{(k)})\|^2}{\|\nabla f(x^{(k-1)})\|^2}$$

Polak-Ribiere: 
$$\beta_k = \frac{\nabla f(x^{(k)})^T y^{(k)}}{\|\nabla f(x^{(k-1)})\|^2}$$

Hestenes-Stiefel: 
$$\beta_k = \frac{\nabla f(x^{(k)})^T y^{(k)}}{d^{(k-1)T} y^{(k)}}$$

där  $y^{(k)} = \nabla f(x^{(k)}) - \nabla f(x^{(k-1)})$ .

(Vi ska inte gå in på varför.)



# Konjugerade gradientmetoder

**Konjugerade gradientmetoder:** Sätt  $d^{(k)} = -\nabla f(x^{(k)}) + \beta_k d^{(k-1)}$ .

Formler för att beräkna  $\beta$ :

$$\text{Fletcher-Reeves: } \beta_k = \frac{\|\nabla f(x^{(k)})\|^2}{\|\nabla f(x^{(k-1)})\|^2}$$

$$\text{Polak-Ribiere: } \beta_k = \frac{\nabla f(x^{(k)})^T y^{(k)}}{\|\nabla f(x^{(k-1)})\|^2}$$

$$\text{Hestenes-Stiefel: } \beta_k = \frac{\nabla f(x^{(k)})^T y^{(k)}}{d^{(k-1)T} y^{(k)}}$$

där  $y^{(k)} = \nabla f(x^{(k)}) - \nabla f(x^{(k-1)})$ .

(Vi ska inte gå in på varför.)

Vilken är bäst?

# Konjugerade gradientmetoder

**Konjugerade gradientmetoder:** Sätt  $d^{(k)} = -\nabla f(x^{(k)}) + \beta_k d^{(k-1)}$ .

Formler för att beräkna  $\beta$ :

$$\text{Fletcher-Reeves: } \beta_k = \frac{\|\nabla f(x^{(k)})\|^2}{\|\nabla f(x^{(k-1)})\|^2}$$

$$\text{Polak-Ribiere: } \beta_k = \frac{\nabla f(x^{(k)})^T y^{(k)}}{\|\nabla f(x^{(k-1)})\|^2}$$

$$\text{Hestenes-Stiefel: } \beta_k = \frac{\nabla f(x^{(k)})^T y^{(k)}}{d^{(k-1)T} y^{(k)}}$$

där  $y^{(k)} = \nabla f(x^{(k)}) - \nabla f(x^{(k-1)})$ .

(Vi ska inte gå in på varför.)

Vilken är bäst? Det beror på  $f(x)$ .

# Konjugerade gradientmetoder

**Konjugerade gradientmetoder:** Sätt  $d^{(k)} = -\nabla f(x^{(k)}) + \beta_k d^{(k-1)}$ .

Formler för att beräkna  $\beta$ :

$$\text{Fletcher-Reeves: } \beta_k = \frac{\|\nabla f(x^{(k)})\|^2}{\|\nabla f(x^{(k-1)})\|^2}$$

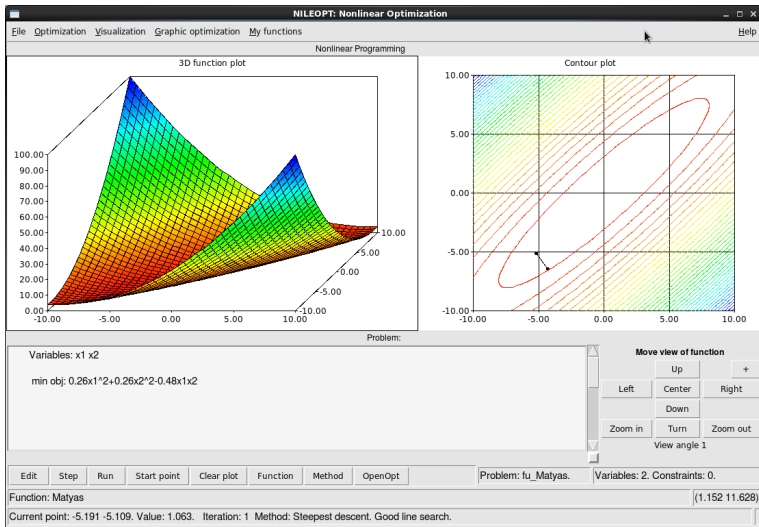
$$\text{Polak-Ribiere: } \beta_k = \frac{\nabla f(x^{(k)})^T y^{(k)}}{\|\nabla f(x^{(k-1)})\|^2}$$

$$\text{Hestenes-Stiefel: } \beta_k = \frac{\nabla f(x^{(k)})^T y^{(k)}}{d^{(k-1)T} y^{(k)}}$$

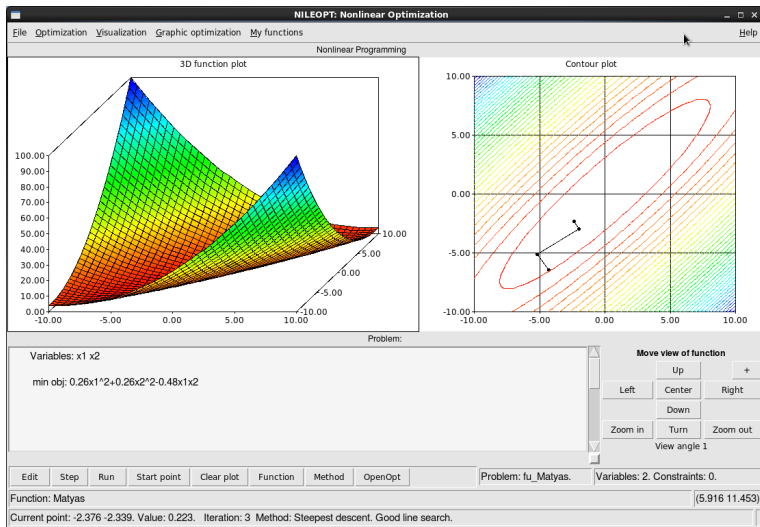
där  $y^{(k)} = \nabla f(x^{(k)}) - \nabla f(x^{(k-1)})$ .

(Vi ska inte gå in på varför.)

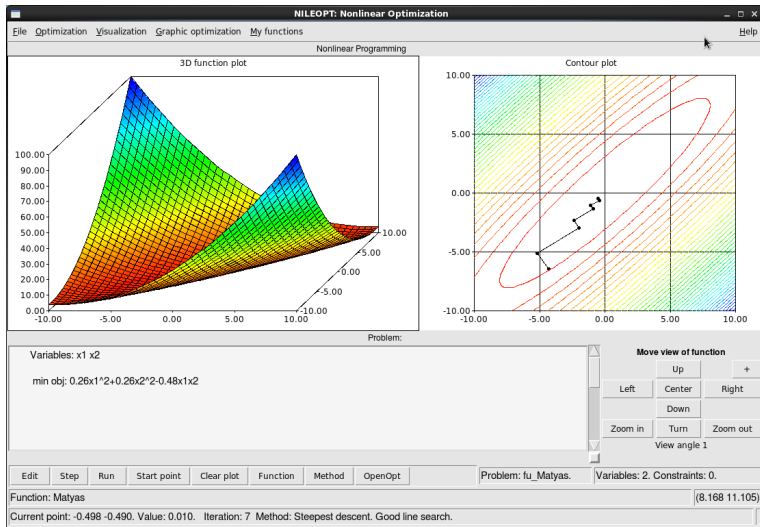
Vilken är bäst? Det beror på  $f(x)$ . Man kan pröva med Nileopt.



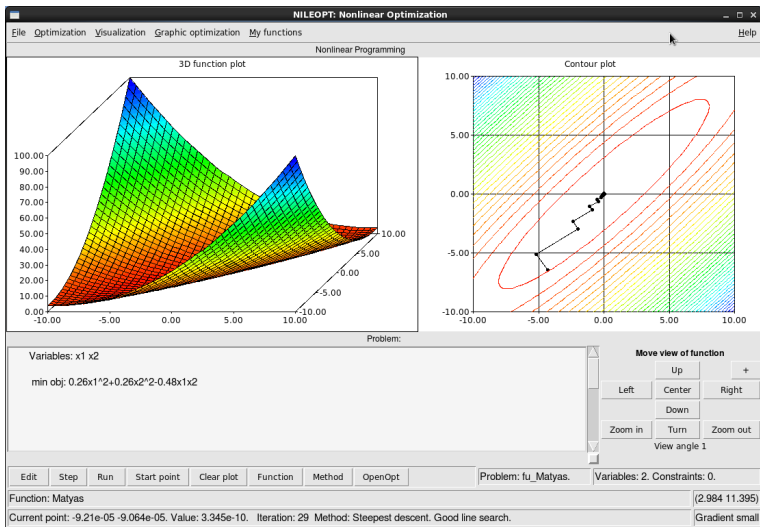
Brantaste lutningsmetoden, iteration 1.



Brantaste lutningsmetoden, iteration 2-3.



Brantaste lutningsmetoden, iteration 4-7.



Brantaste lutningsmetoden, iteration 8-29.

Format för olinjära optimeringsproblem i Nileopt (GMPL,AMPL):



# Nileopt

Format för olinjära optimeringsproblem i Nileopt (GMPL,AMPL):

Problemet

$$\min f(x) = x_1^2 + 2x_2^2 + x_1x_2 + 2x_1 - x_2$$

# Nileopt

Format för olinjära optimeringsproblem i Nileopt (GMPL,AMPL):

Problemet

$$\min f(x) = x_1^2 + 2x_2^2 + x_1x_2 + 2x_1 - x_2$$

skrivs in i Nileopt som

```
var x1;
```

```
var x2;
```

```
min obj: x1^2+2*x2^2+x1*x2+2*x1-1*x2;
```

# Nileopt

Format för olinjära optimeringsproblem i Nileopt (GMPL,AMPL):

Problemet

$$\min f(x) = x_1^2 + 2x_2^2 + x_1x_2 + 2x_1 - x_2$$

skrivs in i Nileopt som

```
var x1;  
var x2;  
min obj: x1^2+2*x2^2+x1*x2+2*x1-1*x2;
```

För att få fram "upphöjt till" (^) krävs ett extra mellanslag.

# Nileopt: Format

Problemet

$$\min f(x) = 2x_1^2 - 2x_1x_2 + 3x_2$$

$$x_1^2 + x_2^2 \leq 1$$

då  $x_1 + x_1x_2 \leq 1$

$$x_1 + x_2 \leq 2$$

## Nileopt: Format

Problemet

$$\min f(x) = 2x_1^2 - 2x_1x_2 + 3x_2$$

$$x_1^2 + x_2^2 \leq 1$$

$$\text{då } x_1 + x_1x_2 \leq 1$$

$$x_1 + x_2 \leq 2$$

kan skrivas in i Nileopt som

```
var x1;
```

```
var x2;
```

```
minimize obj:
```

```
2*x1^2 - 2*x1*x2 + 3*x2;
```

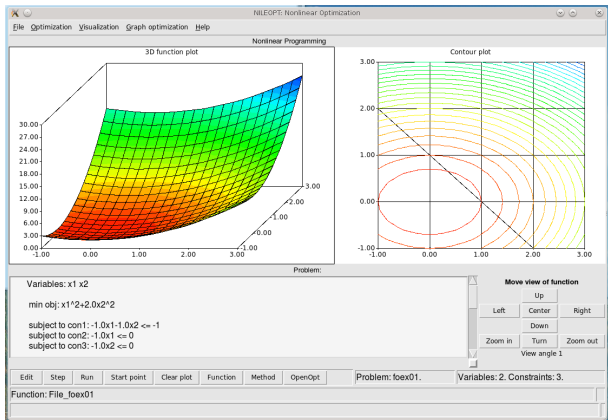
```
subject to con1: x1^2 + x2^2 <= 1;
```

```
s.t. con2: x1 + x1*x2 <= 1;
```

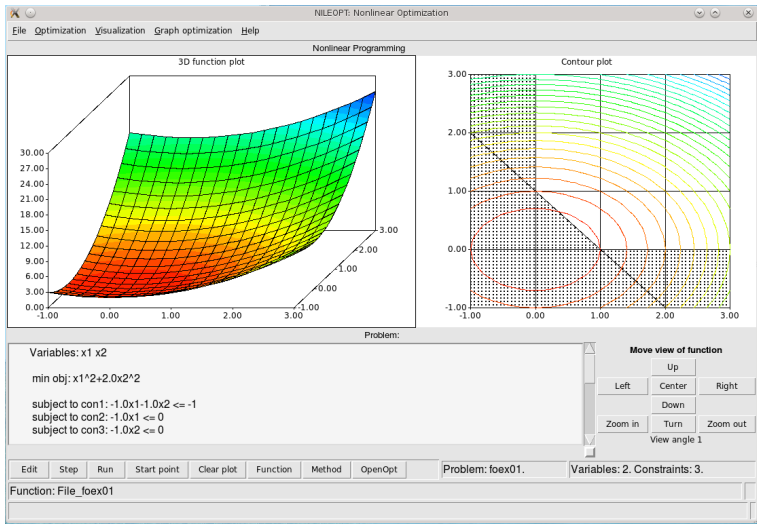
```
st con3: x1+x2<=2;
```

# Nileopt med bivillkor

$$\begin{aligned} \min \quad & f(x) = x_1^2 + 2x_2^2 \\ \text{då} \quad & x_1 + x_2 \geq 1 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{aligned}$$



# Nileopt med bivillkor



# Laboration 3

I laboration 3 ska ni:



# Laboration 3

I laboration 3 ska ni:

- Skriva in olinjära optimeringsproblem enligt reglerna.

## Laboration 3

I laboration 3 ska ni:

- Skriva in olinjära optimeringsproblem enligt reglerna.
- Lösa olinjära problem med bivillkor med färdig programvara (CobyLa).

## Laboration 3

I laboration 3 ska ni:

- Skriva in olinjära optimeringsproblem enligt reglerna.
- Lösa olinjära problem med bivillkor med färdig programvara (CobyLa).
- Illustrera KKT-villkoren grafiskt.

# Laboration 3

I laboration 3 ska ni:

- Skriva in olinjära optimeringsproblem enligt reglerna.
- Lösa olinjära problem med bivillkor med färdig programvara (CobyLa).
- Illustrera KKT-villkoren grafiskt.
- Testa olika metoder för olinjära problem utan bivillkor på olika funktioner.

# Laboration 3

I laboration 3 ska ni:

- Skriva in olinjära optimeringsproblem enligt reglerna.
- Lösa olinjära problem med bivillkor med färdig programvara (CobyLa).
- Illustrera KKT-villkoren grafiskt.
- Testa olika metoder för olinjära problem utan bivillkor på olika funktioner.
- Intressanta frågor:

# Laboration 3

I laboration 3 ska ni:

- Skriva in olinjära optimeringsproblem enligt reglerna.
- Lösa olinjära problem med bivillkor med färdig programvara (CobyLa).
- Illustrera KKT-villkoren grafiskt.
- Testa olika metoder för olinjära problem utan bivillkor på olika funktioner.
- Intressanta frågor:

Hur mycket sämre är brantaste lutningsmetoden än Newtons metod?

# Laboration 3

I laboration 3 ska ni:

- Skriva in olinjära optimeringsproblem enligt reglerna.
- Lösa olinjära problem med bivillkor med färdig programvara (CobyLa).
- Illustrera KKT-villkoren grafiskt.
- Testa olika metoder för olinjära problem utan bivillkor på olika funktioner.
- Intressanta frågor:

Hur mycket sämre är brantaste lutningsmetoden än Newtons metod?

Är sicksackning verkligen ett problem för brantaste lutningsmetoden?

# Laboration 3

I laboration 3 ska ni:

- Skriva in olinjära optimeringsproblem enligt reglerna.
- Lösa olinjära problem med bivillkor med färdig programvara (CobyLa).
- Illustrera KKT-villkoren grafiskt.
- Testa olika metoder för olinjära problem utan bivillkor på olika funktioner.
- Intressanta frågor:

Hur mycket sämre är brantaste lutningsmetoden än Newtons metod?

Är sicksackning verkligen ett problem för brantaste lutningsmetoden?

Är konjugerade gradientmetoder bättre än brantaste lutningsmetoden?



# Laboration 3

I laboration 3 ska ni:

- Skriva in olinjära optimeringsproblem enligt reglerna.
- Lösa olinjära problem med bivillkor med färdig programvara (Cobyli).
- Illustrera KKT-villkoren grafiskt.
- Testa olika metoder för olinjära problem utan bivillkor på olika funktioner.
- Intressanta frågor:

Hur mycket sämre är brantaste lutningsmetoden än Newtons metod?

Är sicksackning verkligen ett problem för brantaste lutningsmetoden?

Är konjugerade gradientmetoder bättre än brantaste lutningsmetoden?

Är kvasi-Newtonmetoder bättre än brantaste lutningsmetoden?

## Laboration 3

I laboration 3 ska ni:

- Skriva in olinjära optimeringsproblem enligt reglerna.
- Lösa olinjära problem med bivillkor med färdig programvara (Cobyli).
- Illustrera KKT-villkoren grafiskt.
- Testa olika metoder för olinjära problem utan bivillkor på olika funktioner.
- Intressanta frågor:

Hur mycket sämre är brantaste lutningsmetoden än Newtons metod?

Är sicksackning verkligen ett problem för brantaste lutningsmetoden?

Är konjugerade gradientmetoder bättre än brantaste lutningsmetoden?

Är kvasi-Newtonmetoder bättre än brantaste lutningsmetoden?

Är kvasi-Newtonmetoder sämre än Newtons metod?

## Laboration 3

I laboration 3 ska ni:

- Skriva in olinjära optimeringsproblem enligt reglerna.
- Lösa olinjära problem med bivillkor med färdig programvara (Cobyli).
- Illustrera KKT-villkoren grafiskt.
- Testa olika metoder för olinjära problem utan bivillkor på olika funktioner.
- Intressanta frågor:

Hur mycket sämre är brantaste lutningsmetoden än Newtons metod?

Är sicksackning verkligen ett problem för brantaste lutningsmetoden?

Är konjugerade gradientmetoder bättre än brantaste lutningsmetoden?

Är kvasi-Newtonmetoder bättre än brantaste lutningsmetoden?

Är kvasi-Newtonmetoder sämre än Newtons metod?

Är icke-konvexa problem svårare än konvexa?