Optimering med bivillkor

Flervariabelanalys

Linköpings Universitet

Översikt

$M \subset \mathbb{R}^n$ och reellvärda funktioner $f, g_1, g_2, \ldots, g_k$ som är kontinuerliga på M och av klass $\mathcal{C}^1$ i M^o.

$$D = \{\bar{x} \in M : g_1(\bar{x}) = c_1, g_2(\bar{x}) = c_2, \ldots, g_k(\bar{x}) = c_k\}.$$

Ett bivillkor i två dimensioner

Låt

$$D = \{(x, y) \in M : g(x, y) = c\},$$

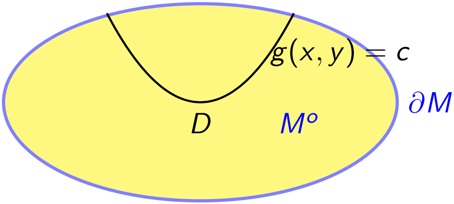
och antag att $g : M \to \mathbb{R}$, $f : M \to \mathbb{R}$ är av klass $\mathcal{C}^1$ i M^o.

Ett bivillkor i två dimensioner

Låt

$$D = \{(x, y) \in M : g(x, y) = c\},$$

och antag att $g : M \to \mathbb{R}$, $f : M \to \mathbb{R}$ är av klass $\mathcal{C}^1$ i M^o. Vi har ett exempel på en sådan mängd D nedan. Här är M en ellipsskiva, och mängden D är snittet mellan denna skiva och nivåkurvan $g(x, y) = c$ (i detta fall $x^2 - y = 0$).



Sats

Låt

$$D = \{(x, y) \in M : g(x, y) = c\},$$

och antag att $g : M \to \mathbb{R}$, $f : M \to \mathbb{R}$ är av klass $\mathcal{C}^1$ i M^o. Då gäller att om $f : D \to \mathbb{R}$ har en lokal extrempunkt över D i $(a, b) \in M^o$ så är $\nabla f(a, b), \nabla g(a, b)$ linjärt beroende.

Sats

Låt

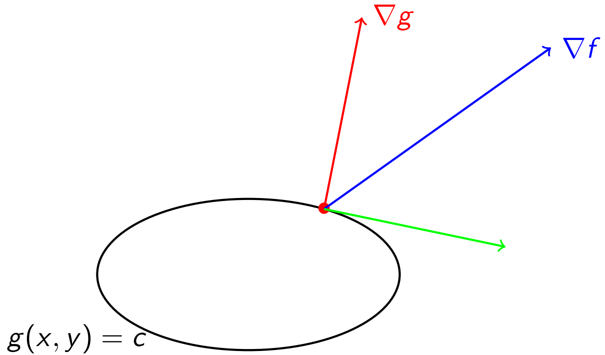
$$D = \{(x, y) \in M : g(x, y) = c\},$$

och antag att $g : M \to \mathbb{R}$, $f : M \to \mathbb{R}$ är av klass $\mathcal{C}^1$ i M^o. Då gäller att om $f : D \to \mathbb{R}$ har en lokal extrempunkt över D i $(a, b) \in M^o$ så är $\nabla f(a, b), \nabla g(a, b)$ linjärt beroende.

Villkoret att vektorerna $\nabla f(a, b), \nabla g(a, b)$ är linjärt beroende kan formuleras som att antingen $\nabla g(a, b) = 0$ eller så finns det λ sådant att $\nabla f(a, b) = \lambda \nabla g(a, b)$. Detta kallas **Lagranges multiplikatormetod**, och λ kallas en **Lagrangemultiplikator**.

Villkoret att vektorerna $\nabla f(a,b), \nabla g(a,b)$ är linjärt beroende kan formuleras som att antingen $\nabla g(a,b) = 0$ eller så finns det λ sådant att $\nabla f(a,b) = \lambda \nabla g(a,b)$. Detta kallas **Lagranges multiplikatormetod**, och λ kallas en **Lagrangemultiplikator**. Det är även ekvivalent med att determinanten av matrisen vi får om vi sätter in de två vektorerna i raderna i en matris (som i detta fall blir en 2×2-matris) är noll.

Motivering



Ett bivillkor i tre dimensioner

Låt

$$D = \{(x, y, z) \in M : g(x, y, z) = k\},$$

och antag att $g : M \to \mathbb{R}$, $f : M \to \mathbb{R}$ är av klass $\mathcal{C}^1$ i M^o. Då gäller att om $f : D \to \mathbb{R}$ har en lokal extrempunkt över D i $(a, b, c) \in M^o$ då är $\nabla f(a, b, c), \nabla g(a, b, c)$ linjärt beroende.

Villkoret att vektorerna $\nabla f(a,b,c), \nabla g(a,b,c)$ är linjärt beroende kan formuleras som att antingen $\nabla g(a,b,c) = 0$ eller så finns det λ sådant att $\nabla f(a,b,c) = \lambda \nabla g(a,b,c)$. Detta kallas **Lagranges multiplikatormetod**, och λ kallas en **Lagrangemultiplikator**.

Villkoret att vektorerna $\nabla f(a, b, c), \nabla g(a, b, c)$ är linjärt beroende kan formuleras som att antingen $\nabla g(a, b, c) = 0$ eller så finns det λ sådant att $\nabla f(a, b, c) = \lambda \nabla g(a, b, c)$. Detta kallas **Lagranges multiplikatormetod**, och λ kallas en **Lagrangemultiplikator**. Det är även ekvivalent med att kryssprodukten mellan de två vektorerna är noll.

Två bivillkor i tre dimensioner

Om vi har

$$D = \{(x, y, z) \in M : g(x, y, z) = k, h(x, y, z) = l\}$$

och antag att $g : M \to \mathbb{R}$,$h : M \to \mathbb{R}$ och $f : M \to \mathbb{R}$ är av klass $\mathcal{C}^1$ i M^o. Då gäller att om $f : D \to \mathbb{R}$ har en lokal extrempunkt över D i $(a, b, c) \in M^o$ då är vektorerna $\nabla f(a, b, c), \nabla g(a, b, c), \nabla h(a, b, c)$ linjärt beroende.

Villkoret att vektorerna $\nabla f(a, b, c), \nabla g(a, b, c), \nabla h(a, b, c)$ är linjärt beroende kan formuleras som att antingen $\nabla g(a, b, c), \nabla h(a, b, c)$ är linjärt beroende, eller så finns λ_1, λ_2 sådana att $\nabla f(a, b, c) = \lambda_1 \nabla g(a, b, c) + \lambda_2 \nabla h(a, b, c)$. Detta kallas **Lagranges multiplikatormetod**, och λ_1, λ_2 kallas **Lagrangemultiplikatorer**.

Det är även ekvivalent med att determinanten av matrisen vi får om vi sätter in de tre vektorerna i raderna i en matris (som i detta fall blir en 3×3-matris) är noll.

Motivering

Om $\nabla g(a, b, c)$ och $\nabla h(a, b, c)$ inte är linjärt beroende, då säger implicita funktionssatsen att mängden där $g(x, y, z) = k$ och $h(x, y, z) = l$ lokalt kring (a, b, c) är en parameterkurva $\bar{x}(t)$ där $(a, b, c) = \bar{x}(0)$.

Motivering

Om $\nabla g(a,b,c)$ och $\nabla h(a,b,c)$ inte är linjärt beroende, då säger implicita funktionssatsen att mängden där $g(x,y,z)=k$ och $h(x,y,z)=l$ lokalt kring (a,b,c) är en parameterkurva $\bar{x}(t)$ där $(a,b,c)=\bar{x}(0)$.
Om f har en lokal extrempunkt över denna i $\bar{x}(0)$ då gäller att

$$\frac{d}{dt}f(\bar{x}(t))=\nabla f(\bar{x}(t))\bullet \bar{x}'(t)$$

är noll i $t=0$.

Motivering

Om $\nabla g(a,b,c)$ och $\nabla h(a,b,c)$ inte är linjärt beroende, då säger implicita funktionssatsen att mängden där $g(x,y,z)=k$ och $h(x,y,z)=l$ lokalt kring (a,b,c) är en parameterkurva $\bar{x}(t)$ där $(a,b,c)=\bar{x}(0)$.
Om f har en lokal extrempunkt över denna i $\bar{x}(0)$ då gäller att

$$\frac{d}{dt}f(\bar{x}(t)) = \nabla f(\bar{x}(t)) \bullet \bar{x}'(t)$$

är noll i $t=0$.
Eftersom $\bar{x}'(0)$ pekar i tangentriktningen till kurvan säger detta att $\nabla f(a,b,c)$ är ortogonal mot tangentriktningen.

Motivering

Om $\nabla g(a, b, c)$ och $\nabla h(a, b, c)$ inte är linjärt beroende, då säger implicita funktionssatsen att mängden där $g(x, y, z) = k$ och $h(x, y, z) = l$ lokalt kring (a, b, c) är en parameterkurva $\bar{x}(t)$ där $(a, b, c) = \bar{x}(0)$.
Om f har en lokal extrempunkt över denna i $\bar{x}(0)$ då gäller att

$$\frac{d}{dt} f(\bar{x}(t)) = \nabla f(\bar{x}(t)) \bullet \bar{x}'(t)$$

är noll i $t = 0$.
Eftersom $\bar{x}'(0)$ pekar i tangentriktningen till kurvan säger detta att $\nabla f(a, b, c)$ är ortogonal mot tangentriktningen.
Men detta gäller även för båda vektorerna $\nabla g(a, b, c)$ och $\nabla h(a, b, c)$ också.

Motivering

Om $\nabla g(a,b,c)$ och $\nabla h(a,b,c)$ inte är linjärt beroende, då säger implicita funktionssatsen att mängden där $g(x,y,z)=k$ och $h(x,y,z)=l$ lokalt kring (a,b,c) är en parameterkurva $\bar{x}(t)$ där $(a,b,c)=\bar{x}(0)$.

Om f har en lokal extrempunkt över denna i $\bar{x}(0)$ då gäller att

$$\frac{d}{dt}f(\bar{x}(t))=\nabla f(\bar{x}(t))\bullet\bar{x}'(t)$$

är noll i $t=0$.

Eftersom $\bar{x}'(0)$ pekar i tangentriktningen till kurvan säger detta att $\nabla f(a,b,c)$ är ortogonal mot tangentriktningen.

Men detta gäller även för båda vektorerna $\nabla g(a,b,c)$ och $\nabla h(a,b,c)$ också.

Eftersom dessa är linjärt oberoende betyder det att $\nabla f(a,b,c)$ ligger i det plan dessa två spänner upp, vilket precis betyder att $\nabla f(a,b,c)=\lambda_1\nabla g(a,b,c)+\lambda_2\nabla h(a,b,c)$ för några λ_1,λ_2.