



Föreläsning 10: NLP med likhetsbivillkor och KKT villkoren

Allmänna icke-linjära problem under likhetsbivillkor

$$\begin{aligned} &\text{minimera } f(\mathbf{x}) \\ &\text{då } h_i(\mathbf{x}) = 0, \quad i = 1, \dots, m \end{aligned} \tag{1}$$

Det tillåtna området $\mathcal{F} = \{\mathbf{x} \in \mathbf{R}^n : h_i(\mathbf{x}) = 0, \quad i = 1, \dots, m\}$ är inte konvext såvida inte funktionerna h_i är affina, d.v.s. $h_i(\mathbf{x}) = \mathbf{a}_i^T \mathbf{x} + b_i$.

Vi antar att $n > m$.

Vi behöver följande tekniska antagande:

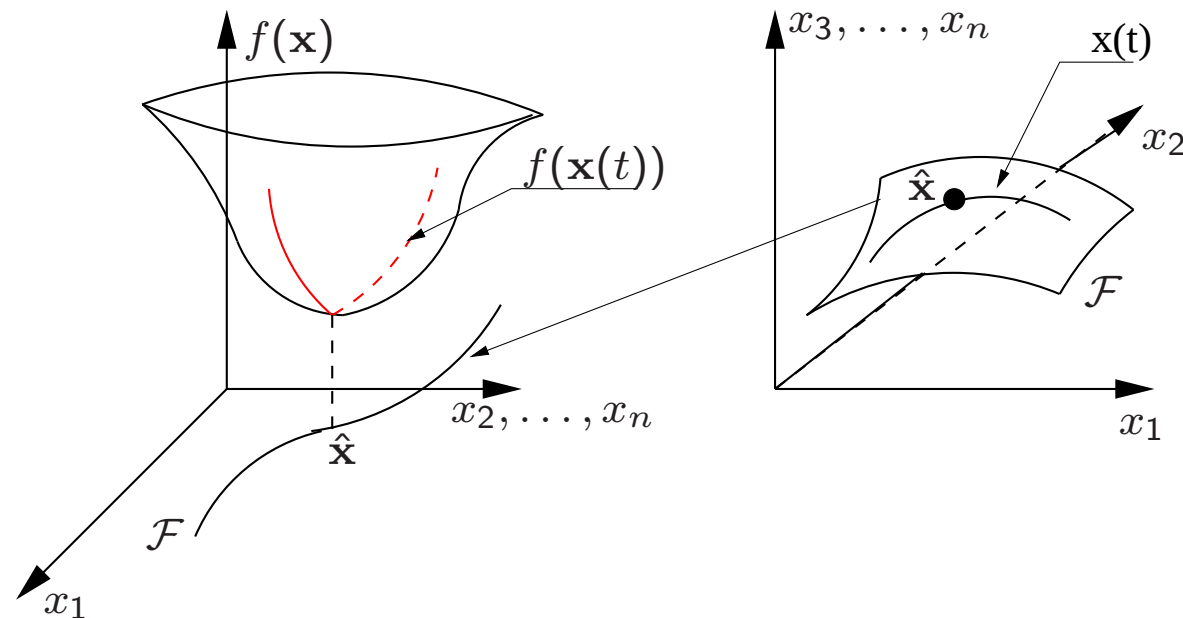
Definition 1. En tillåten lösning $\mathbf{x} \in \mathcal{F}$ är en reguljär punkt till (1) om $\nabla h_i(\mathbf{x}), \quad i = 1, \dots, m$ är linjärt oberoende.

Sats 1 (KKT för problem med likhetsbivillkor). *Antag att $\hat{\mathbf{x}} \in \mathcal{F}$ är en reguljär punkt och en lokal optimallösning till (1). Då finns $\hat{\mathbf{u}} \in \mathbf{R}^m$ så att*

$$(1) \quad \nabla f(\hat{\mathbf{x}}) + \sum_{i=1}^m \hat{u}_i \nabla h_i(\hat{\mathbf{x}}) = \mathbf{0}^T,$$

$$(2) \quad h_i(\hat{\mathbf{x}}) = 0, \quad i = 1, \dots, m.$$

Bevisskiss: Låt $\mathbf{x}(t)$ vara en godtycklig parametriserad kurva i tillåtna mängden $\mathcal{F} = \{\mathbf{x} \in \mathbf{R}^n : g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m\}$ sådan att $\mathbf{x}(0) = \hat{\mathbf{x}}$. Figuren på nästa sida illustrerar hur denna parametriserade kurva avbildas på en kurva $f(\mathbf{x}(t))$ på målfunktionsytan. Den tillåtna mängden $\mathcal{F}$ är i allmänhet av högre dimension än 1 vilket illustreras med hjälp av högra figuren.



Eftersom $\mathbf{x}(0) = \hat{\mathbf{x}}$ är en lokal optimal lösning gäller att

$$\frac{d}{dt} f(\mathbf{x}(t))|_{t=0} = \nabla f(\hat{\mathbf{x}}) \cdot \mathbf{x}'(0) = 0$$

Vidare gäller att $\mathbf{x}(t) \in \mathcal{F}$ vilket medför att

$$h_i(\mathbf{x}(t)) = 0, \quad i = 1, \dots, m, \quad \forall t \in (-\epsilon, \epsilon)$$

för något $\epsilon > 0$.

Detta innebär att

$$\frac{d}{dt}h_i(\mathbf{x}(t))|_{t=0} = \nabla h_i(\hat{\mathbf{x}}) \cdot \mathbf{x}'(0) = 0, \quad i = 1, \dots, m$$

vilket i sin tur innebär att $\mathbf{x}'(0) \in \mathcal{N}(\mathbf{A})$, där

$$\mathbf{A} = \begin{bmatrix} \nabla h_1(\hat{\mathbf{x}}) \\ \vdots \\ \nabla h_m(\hat{\mathbf{x}}) \end{bmatrix}$$

Omvänt kan man med hjälp av implicita funktionssatsen visa att om $\mathbf{p} \in \mathcal{N}(\mathbf{A})$ så finns en parametriserad kurva $\mathbf{x}(t) \in \mathcal{F}$ med $\mathbf{x}(0) = \hat{\mathbf{x}}$ och $\mathbf{x}'(0) = \mathbf{p}$.

Sammantaget betyder ovanstående resonemang att

$$\begin{aligned}\nabla f(\hat{\mathbf{x}})\mathbf{p} &= 0, \quad \forall \mathbf{p} \in \mathcal{N}(\mathbf{A}) \\ \Leftrightarrow \quad \nabla f(\hat{\mathbf{x}})^\top &\in \mathcal{N}(\mathbf{A})^\perp = \mathcal{R}(\mathbf{A}^\top) \\ \Leftrightarrow \quad \nabla f(\hat{\mathbf{x}})^\top &= \mathbf{A}^\top \hat{\mathbf{v}},\end{aligned}$$

för något $\hat{\mathbf{v}} \in \mathbf{R}^m$.

Om vi låter $\hat{\mathbf{u}} = -\hat{\mathbf{v}} \in \mathbf{R}^m$ kan det sista uttrycket skrivas

$$\nabla f(\hat{\mathbf{x}}) + \sum_{i=1}^m \hat{u}_i \nabla h_i(\hat{\mathbf{x}}) = \mathbf{0}^\top$$

vilket skulle visas.

Läsanvisningar

Kapitel 18 och 19 i optimeringskompendiet.

- Ickelinjär optimering, grå häftet, sidorna 18-20.