Optimierung

Vorlesung 4 (Erweiterung)

Erklärung Dualität

Optimierung mit Nebenbedingungen: Äquivalente Formulierungen

- Formulierung 1: $\min_x f(x)$ subject to $c_i(x) \geq 0, \; i \in \mathcal{I}$
- Dazugehörige Lagrangefunktion: $\mathcal{L}(x, \lambda) = f(x) - \sum_i \lambda_i c_i(x)$

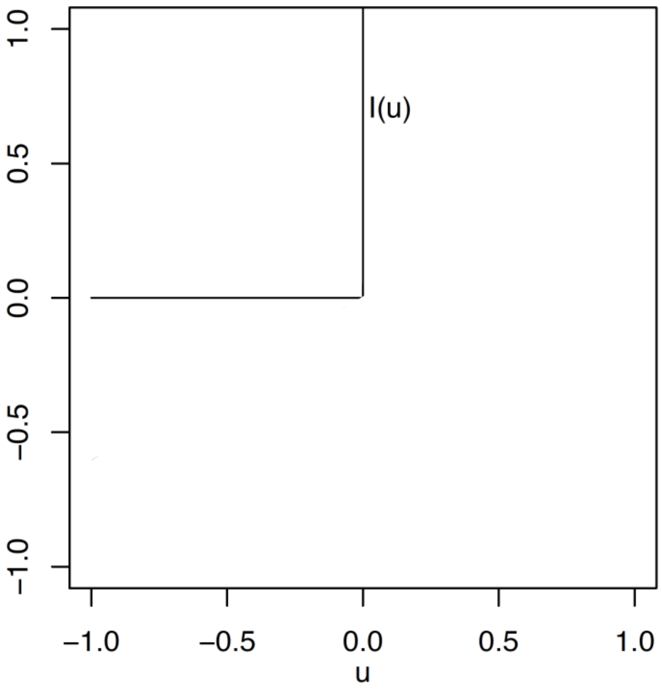
- Äquivalente Formulierung: $\min_x f(x)$ subject to $c_i(x) \leq 0, \; i \in \mathcal{I}$
- Dazugehörige Lagrangefunktion: $\mathcal{L}(x, \lambda) = f(x) + \sum_i \lambda_i c_i(x)$

Step Funktion

- Lagrangefunktion: $\mathcal{L}(x,\lambda) = f(x) + \sum_i \lambda_i c_i(x)$
- Eigentlich wollen wir nicht $\mathcal{L}(x,\lambda)$ minimieren, sondern folgende Funktion:

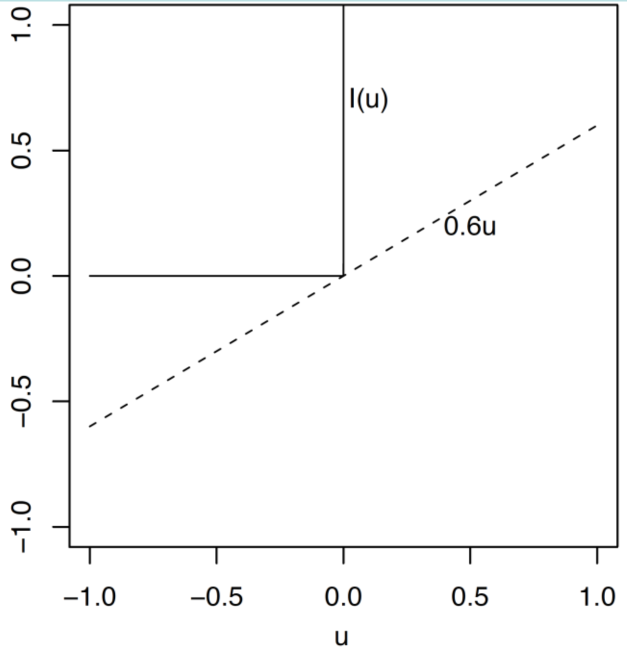
$$J(x) = \begin{cases} f(x) & \text{if } c_i(x) \leq 0 \\ \infty & \text{else} \end{cases}$$
$$= f(x) + \sum_i I[c_i(x))]$$

$$\text{with } I[u] = \begin{cases} 0 & \text{if } u \leq 0 \\ \infty & \text{else} \end{cases}$$



- Wir suchen also: $\min_x J(x)$
- Problem: $I[u]$ nicht differenzierbar

Step Funktion und Lagrange Funktion



- Daher: Approximiere $I[u]$ durch lineare Funktion $\lambda_i \cdot u$

- Aus $J(x) = f(x) + \sum_i I[c_i(x))]$ wird dann:

$$\mathcal{L}(x, \lambda) = f(x) + \sum_i \lambda_i c_i(x)$$

- Aus dem Bild sieht man sofort dass die Lagrange-Funktion eine untere Schranke ist:

$$\forall \lambda \geq 0 : \mathcal{L}(x, \lambda) \leq J(x)$$

- Gegeben ein festes x, was ist das "beste" λ_i?

$$c_i(x) \leq 0 \quad \Rightarrow \quad \lambda_i = 0 \quad (\text{Bem.: } \lambda > 0 \text{ verkleinert da } c_i(x) \leq 0)$$

$$c_i(x) > 0 \quad \Rightarrow \quad \lambda_i \to \infty$$

- Daher gilt (etwas überraschend?): $\max_\lambda \mathcal{L}(x, \lambda) = J(x)$

Komplementaritätsbedingung

- Da wir an $\min_x J(x)$ interessiert sind, suchen wir nach

$$\min_x \max_\lambda \mathcal{L}(x,\lambda) = \min_x J(x)$$

- Das ist aber vielleicht schwer, daher schauen wir mal auf

$$\max_\lambda \min_x \mathcal{L}(x,\lambda)$$

- Daraus leiten wir das duale Optimierungsproblem

$$\max_\lambda g(\lambda) \quad \text{wobei} \quad g(\lambda) = \min_x \mathcal{L}(x,\lambda)$$

- Aus $\forall\lambda \geq 0 : \mathcal{L}(x,\lambda) \leq J(x)$ folgt

$$\min_x \mathcal{L}(x,\lambda) = g(\lambda) \leq \min_x J(x)$$

- Damit ist $\max_\lambda g(\lambda)$ das Problem des Findens der besten unteren Schranke

- Wird das duale Problem genannt und ist einfacher, da konkav (siehe später)

Dualität

- Durch die Nebenbedingungen kamen neben den primalen Variablen x die Lagrange Multiplikatoren λ als Variablen ins Spiel

$$\mathcal{L}(x,\lambda) = f(x) + \sum_i \lambda_i c_i(x), \quad \lambda \geq 0$$

- Sie geben an wie stark die Funktion gegen die Bedingungen drückt.

$$\nabla f(x) = \lambda^\top \nabla c(x), \quad \lambda \geq 0$$

- Alternativ können wir $\mathcal{L}(x,\lambda)$ als zu minimierende Funktion <u>ohne</u> Nebenbedingungen betrachten, indem wir λ festhalten

$$q(\lambda) = \min_x \mathcal{L}(x,\lambda) \quad *$$

- In vielen Fällen erhält man $-\infty$ für einige Werte von λ. Daher schränken wir die Definitionsmenge von $q(\lambda)$ entsprechend ein:

$$\mathcal{D} = \{\lambda | q(\lambda) > -\infty\}$$

- **Anmerkung**: 1.) Wir betrachten nur Ungleichheitsbed.

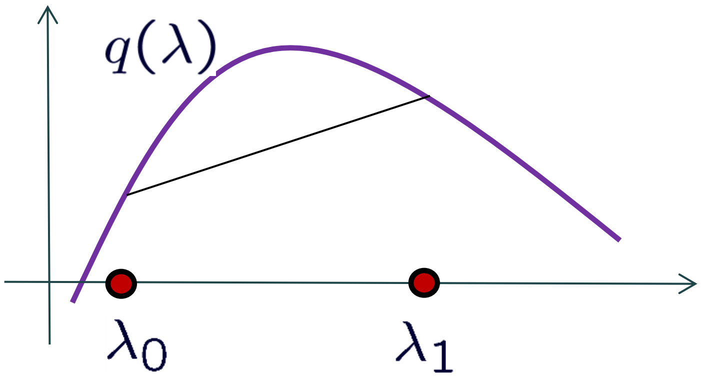
  *.) eigentlich sollten wir inf statt min betrachten

Dualität (Konkav)

- Die Funktion

$$q(\lambda) = \min_x \mathcal{L}(x, \lambda)$$

ist konkav.



- Beweis Schritt 1: Wir zeigen

$$\mathcal{L}(x, (1-\alpha)\lambda_0 + \alpha\lambda_1) = (1-\alpha)\mathcal{L}(x, \lambda_0) + \alpha\mathcal{L}(x, \lambda_1)$$

- Teilbeweis aus def. Lagrange:

$$\begin{aligned} \mathcal{L}(x, (1-\alpha)\lambda_0 + \alpha\lambda_1) &= f(x) + (1-\alpha)\lambda_0 c(x) + \alpha\lambda_1 c(x) \\ &= (1-\alpha)f(x) + \alpha f(x) + (1-\alpha)\lambda_0 c(x) + \alpha\lambda_1 c(x) \\ &= (1-\alpha)\mathcal{L}(x, \lambda) + \alpha\mathcal{L}(x, \lambda) \end{aligned}$$

Dualität (Konkav, cont.)

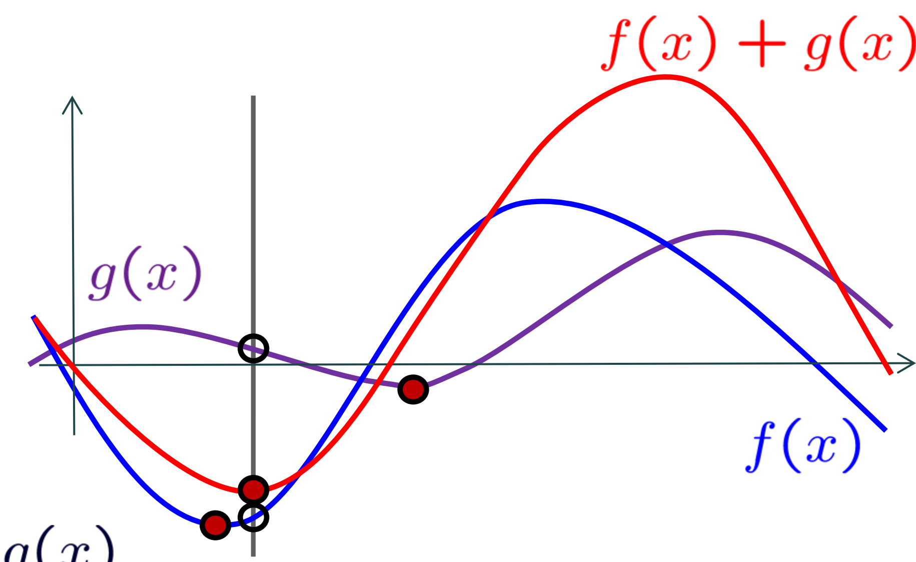
- Damit haben wir:

$$\mathcal{L}(x,(1-\alpha)\lambda_0+\alpha\lambda_1) = (1-\alpha)\mathcal{L}(x,\lambda_0)+\alpha\mathcal{L}(x,\lambda_1)$$

- Beweis (fort.)
- Das Infimum einer Summe ist größer gleich der Summe von Infima, daher:

$$q((1-\alpha)\lambda_0 + \alpha\lambda_1) \geq (1-\alpha)q(\lambda_0) + \alpha q(\lambda_1)$$

- Damit ist Konkavität bewiesen



$$\inf(f(x)+g(x)) \geq \inf f(x) + \inf g(x)$$

Duales Problem (Lagrange Dualität)

- Aus dem primalen Optimierungsproblem

$$\min_x f(x), \qquad c_i(x) \leq 0$$

können wir also ein **duales Optimierungsproblem** ableiten

$$\max_\lambda \min_x \mathcal{L}(x, \lambda), \qquad \lambda \geq 0$$

- In manchen Fällen lässt sich das duale Problem leichter lösen als das primale Problem.

- Sogenannte Primal-Dual-Verfahren optimieren gleichzeitig das primale und das duale Problem.

- Es gibt noch andere Formen der Dualität, z.B. die Fenchel Dualität.

Schwache und starke Dualität

- **Schwache Dualität:**
  Für alle gültigen Lösungen x und λ gilt $q(\lambda) \leq f(x)$

  Beweis: $q(\lambda) = \inf_x f(x) + \lambda^\top c(x) \leq f(x) + \underbrace{\lambda^\top c(x)}_{\leq 0} \leq f(x)$

  d.h. das duale Problem liefert immer eine untere Schranke für das primale Problem

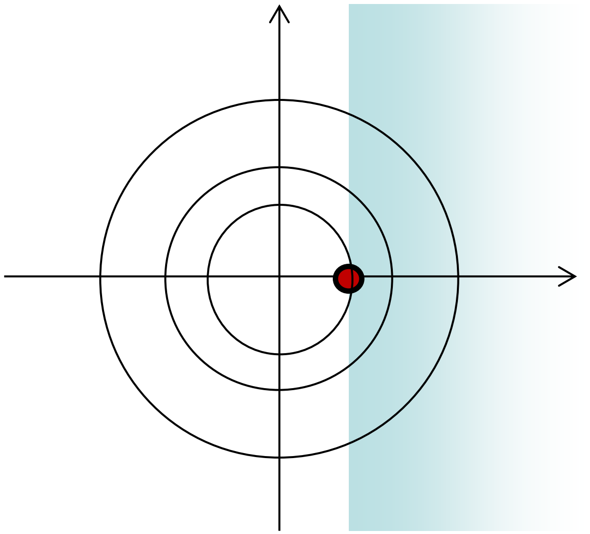
- **Starke Dualität:**
  - $f(x)$ sei konvex und die gültige Menge sei eine konvexe Menge
  - $\hat{\lambda}$ bezeichne das Optimum von $q(\lambda)$ mit $\hat{x} = \mathrm{arginf}_x \mathcal{L}(x, \hat{\lambda})$
  - $\mathcal{L}(x, \hat{\lambda})$ sei <u>streng</u> konvex

  Dann gilt:

  $$q(\hat{\lambda}) = \mathcal{L}(\hat{x}, \hat{\lambda}) = f(\hat{x})$$

  d.h. das Optimum des dualen Problems minimiert auch das primale Problem.

Beispiel



- Konvexes Beispielproblem:

$$\min_{x_1,x_2} \frac{1}{2}(x_1^2 + x_2^2), \qquad 1 - x_1 \leq 0$$

- Lagrange-Funktion:

$$\mathcal{L}(x_1, x_2, \lambda) = \frac{1}{2}(x_1^2 + x_2^2) + \lambda(1 - x_1)$$

- Bedingungen für ein Minimum:

$$x_1 - \lambda = 0, \qquad x_2 = 0,$$

- Damit lässt sich x_1, x_2 in der Lagrange-Funktion eliminieren (geht nicht immer analytisch). Es bleibt das duale Problem

$$\max q(\lambda) = \max -\frac{1}{2}\lambda^2 + \lambda, \qquad \lambda \geq 0$$

mit der Lösung $\lambda = 1$

- Lösung des primalen Problems (starke Dualität): $x_1 = \lambda = 1, \ \ x_2 = 0$

Wichtige Probleme mit starker Dualität

- **Lineare Programme**

$$\min_x w^\top x, \qquad Ax - b \geq 0$$

- **Quadratische Programme**

$$\min_x \frac{1}{2} x^\top Q x + w^\top x, \qquad Ax - b \geq 0$$

mit Q symmetrisch und positiv definit

Zusammenfassung

- Optimalitätsbedingungen für Probleme mit Nebenbedingungen unterscheiden ob eine Bedingung aktiv oder inaktiv ist.
- Beide Fälle lassen sich mithilfe der Lagrange-Funktion ausdrücken.
- Über die Lagrange-Funktion lässt sich auch ein duales Problem herleiten.
- Die Lösung des dualen Problems bildet mindestens eine untere Schranke für das ursprüngliche Problem (schwache Dualität).
- Für konvexe Funktionen mit konvexen Nebenbedingungen führt die Lösung des dualen Problems zur exakten Lösung des ursprünglichen Problems (starke Dualität).

Übungsaufgabe

- Optimierungsproblem:

$$\min_x (x_1 - 3)^2 + (x_2 - 1)^2$$

$$1 - x_1 - x_2 \geq 0 \qquad 1 + x_1 - x_2 \geq 0$$
$$1 - x_1 + x_2 \geq 0 \qquad 1 + x_1 + x_2 \geq 0$$

- Machen Sie eine Skizze des erlaubten Suchraums.

- Berechnen Sie für verschiedene Punkte des Suchraums den Gradienten der Funktion und tragen Sie ihn in die Skizze ein.

- Wo befindet sich das Minimum?

- Wie sehen die Lagrange Multiplikatoren im Optimum aus? Welche Nebenbedingungen sind im Optimum aktiv, welche nicht?