Fundamentals of Linear Algebra and Optimization
Handling Equality Constraints Explicitly

Jean Gallier and Jocelyn Quaintance

CIS Department
University of Pennsylvania

jean@cis.upenn.edu

April 6, 2022

Handling Equality Constraints Explicitly

Sometimes it is desirable to handle equality constraints explicitly.

The only difference is that the Lagrange multipliers associated with *equality constraints* are *not required* to be nonnegative.

Handling Equality Constraints Explicitly

Consider the *Optimization Problem* (P')

$$\begin{aligned} \text{minimize} \quad & J(v) \\ \text{subject to} \quad & \varphi_i(v) \leq 0, \quad i = 1, \ldots, m \\ & \psi_j(v) = 0, \quad j = 1, \ldots, p. \end{aligned}$$

Let us also assume that the functions φ_i are *convex* and that the *equality constraints* ψ_j are *affine* (then they are qualified).

KKT Conditions for Equality Constraints

Theorem. Let $\varphi_i \colon \Omega \to \mathbb{R}$ be m convex inequality constraints and $\psi_j \colon \Omega \to \mathbb{R}$ be p affine equality constraints defined on some open convex subset Ω of a finite-dimensional Euclidean vector space V (more generally, a real Hilbert space V), let $J \colon \Omega \to \mathbb{R}$ be some function, let U be given by

$$U = \{x \in \Omega \mid \varphi_i(x) \leq 0,\ \ \psi_j(x) = 0,\ \ 1 \leq i \leq m,\ 1 \leq j \leq p\},$$

and let $u \in U$ be any point such that the functions φ_i and J are differentiable at u.

KKT Conditions for Equality Constraints

(1) If J has a local minimum at u with respect to U, and if the constraints are *qualified*, then there exist some vectors $\lambda \in \mathbb{R}_+^m$ and $\nu \in \mathbb{R}^p$, such that the *KKT conditions* hold:

$$J_u{}' + \sum_{i=1}^{m} \lambda_i(u)(\varphi_i')_u + \sum_{j=1}^{p} \nu_j(\psi_j')_u = 0,$$

and

$$\sum_{i=1}^{m} \lambda_i(u)\varphi_i(u) = 0, \quad \lambda_i \geq 0, \quad i = 1, \ldots, m.$$

KKT Conditions for Equality Constraints

Equivalently, in terms of gradients, the above conditions are expressed as

$$\nabla J_u + \sum_{i=1}^{m} \lambda_i \nabla(\varphi_i)_u + \sum_{j=1}^{p} \nu_j \nabla(\psi_j)_u = 0$$

and

$$\sum_{i=1}^{m} \lambda_i(u)\varphi_i(u) = 0, \quad \lambda_i \geq 0, \quad i = 1, \ldots, m.$$

KKT Conditions for Equality Constraints

(2) Conversely, if the restriction of J to U is convex and if there exist vectors $\lambda \in \mathbb{R}_+^m$ and $\nu \in \mathbb{R}^p$ such that the KKT conditions hold, then the function J has a (global) minimum at u with respect to U.

Lagrange Dual Function

The Lagrangian $L(v, \lambda, \nu)$ of Problem (P') is defined as

$$L(v, \mu, \nu) = J(v) + \sum_{i=1}^{m} \mu_i \varphi_i(v) + \sum_{j=1}^{p} \nu_i \psi_j(v),$$

where $v \in \Omega$, $\mu \in \mathbb{R}^m_+$, and $\nu \in \mathbb{R}^p$.

Dual Problem

The function $G\colon \mathbb{R}^m_+ \times \mathbb{R}^p \to \mathbb{R}$ given by

$$G(\mu, \nu) = \inf_{v \in \Omega} L(v, \mu, \nu) \quad \mu \in \mathbb{R}^m_+,\ \nu \in \mathbb{R}^p$$

is called the *Lagrange dual function* (or *dual function*), and the *Dual Problem* (D') *is*

$$\begin{aligned} &\text{maximize} \quad G(\mu, \nu) \\ &\text{subject to} \ \ \mu \in \mathbb{R}^m_+,\ \nu \in \mathbb{R}^p. \end{aligned}$$

Observe that the Lagrange multipliers ν *are not restricted* to be nonnegative.

Handling Equality Constraints Explicitly

The duality gap theorem of the last lesson is immediately generalized to Problem $(P^{'})$. We leave the precise statement of this result as an exercise

We now give an example of the Lagrangian dual tecnhique to the *Support Vector Machine* (abbreviated as *SVM*).