AARHUS
UNIVERSITET

Optimization
Problems with Equality Constraints

Henrik Karstoft
@
Aarhus Univerisity, Department of Engineering

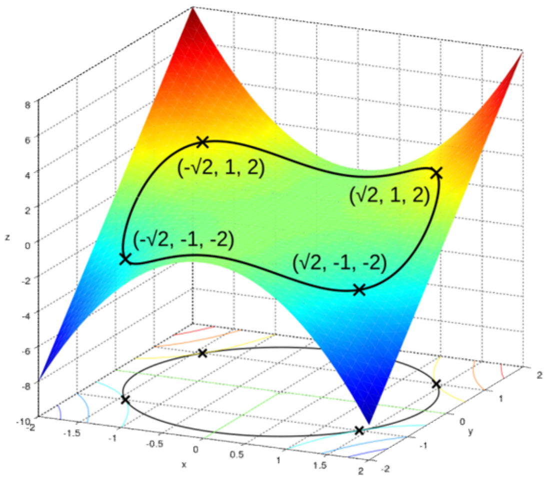
Henrik Karstoft
Aarhus University, Department of Engineering

Contraints optimization

The class of optimization problems we analyze in this chapter is

$$\begin{array}{rl} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \end{array}$$

where $\boldsymbol{x} \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{h} : \mathbb{R}^n \to \mathbb{R}^m$, $\boldsymbol{h} = [h_1, \dots, h_m]^\top$, and $m \leq n$. We assume that the function $\boldsymbol{h}$ is continuously differentiable, that is, $\boldsymbol{h} \in \mathcal{C}^1$.



AARHUS UNIVERSITET

Henrik Karstoft
Aarhus University, Department of Engineering

Contraints optimization, surfaces

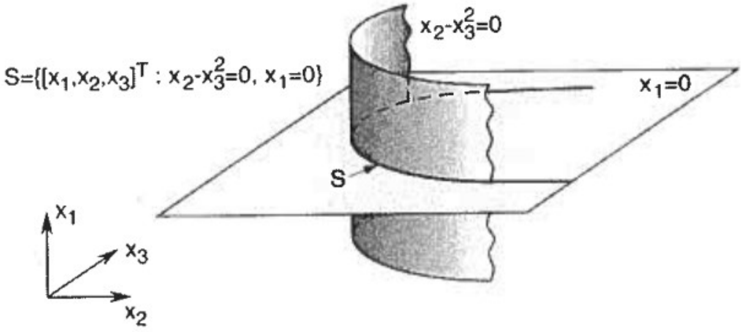
Let $n = 3$ and $m = 2$. Assuming regularity, the feasible set S is a one-dimensional object (i.e., a curve in $\mathbb{R}^3$). For example, let

$$h_1(\boldsymbol{x}) = x_1,$$
$$h_2(\boldsymbol{x}) = x_2 - x_3^2.$$

In this case, $\nabla h_1(\boldsymbol{x}) = [1, 0, 0]^\top$ and $\nabla h_2(\boldsymbol{x}) = [0, 1, -2x_3]^\top$. Hence, the vectors $\nabla h_1(\boldsymbol{x})$ and $\nabla h_2(\boldsymbol{x})$ are linearly independent in $\mathbb{R}^3$. Thus,

$$\dim S = \dim\{\boldsymbol{x} : h_1(\boldsymbol{x}) = 0,\ h_2(\boldsymbol{x}) = 0\} = n - m = 1.$$

See Figure 19.3 for a graphical illustration. ∎



Contraints optimization, tangent space

The *tangent space* at a point x^* on the surface $S = \{x \in \mathbb{R}^n : h(x) = 0\}$ is the set $T(x^*) = \{y : Dh(x^*)y = 0\}$. ∎

Let

$$S = \{x \in \mathbb{R}^3 : h_1(x) = x_1 = 0,\ h_2(x) = x_1 - x_2 = 0\}.$$

Then, S is the x_3-axis in $\mathbb{R}^3$ (see Figure 19.8). We have

$$Dh(x) = \begin{bmatrix} \nabla h_1(x)^\top \\ \nabla h_2(x)^\top \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}.$$

Because ∇h_1 and ∇h_2 are linearly independent when evaluated at any $x \in S$, all the points of S are regular. The tangent space at an arbitrary point of S is

$$\begin{aligned} T(x) &= \{y : \nabla h_1(x)^\top y = 0,\ \nabla h_2(x)^\top y = 0\} \\ &= \left\{ y : \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = 0 \right\} \\ &= \{[0, 0, \alpha]^\top : \alpha \in \mathbb{R}\} \\ &= \text{the } x_3\text{-axis in } \mathbb{R}^3. \end{aligned}$$

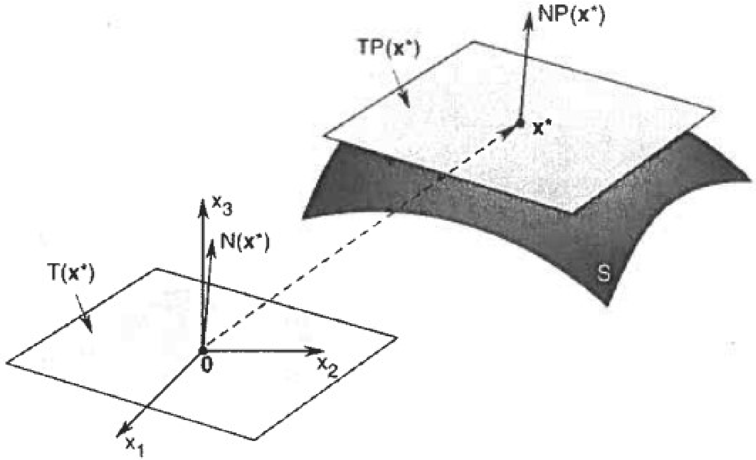
In this example, the tangent space $T(x)$ at any point $x \in S$ is a one-dimensional subspace of $\mathbb{R}^3$. ∎

Henrik Karstoft
Aarhus University, Department of Engineering

Contraints optimization, normal space

Definition The *normal space* $N(\boldsymbol{x}^*)$ at a point $\boldsymbol{x}^*$ on the surface $S = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}\}$ is the set $N(\boldsymbol{x}^*) = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} = D\boldsymbol{h}(\boldsymbol{x}^*)^\top \boldsymbol{z},\ \boldsymbol{z} \in \mathbb{R}^m\}$. ■

Lemma 19.1 *We have* $T(\boldsymbol{x}^*) = N(\boldsymbol{x}^*)^\perp$ *and* $T(\boldsymbol{x}^*)^\perp = N(\boldsymbol{x}^*)$. □



Normal space in $\mathbb{R}^3$.

Lagrange Condition

Lagrange's Theorem for $n = 2$, $m = 1$. *Let the point* x^* *be a minimizer of* $f : \mathbb{R}^2 \to \mathbb{R}$ *subject to the constraint* $h(x) = 0$, $h : \mathbb{R}^2 \to \mathbb{R}$. *Then,* $\nabla f(x^*)$ *and* $\nabla h(x^*)$ *are parallel. That is, if* $\nabla h(x^*) \neq 0$, *then there exists a scalar* λ^* *such that*

$$\nabla f(x^*) + \lambda^* \nabla h(x^*) = 0.$$

□

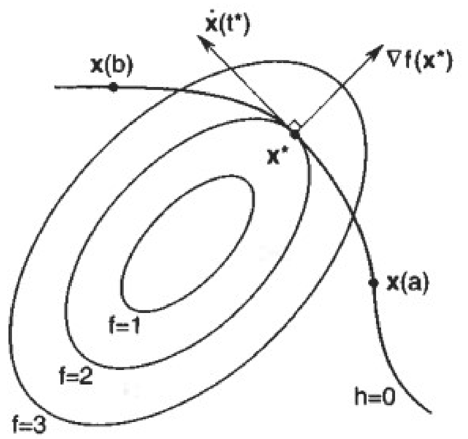


Figure 19.10 The gradient $\nabla f(x^*)$ is orthogonal to the curve $\{x(t)\}$ at the point x^* that is a minimizer of f on the curve.

Lagrange Condition

Lagrange's Theorem. *Let $\boldsymbol{x}^*$ be a local minimizer (or maximizer) of $f : \mathbb{R}^n \to \mathbb{R}$, subject to $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$, $\boldsymbol{h} : \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$. Assume that $\boldsymbol{x}^*$ is a regular point. Then, there exists $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ such that*

$$Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*\top} D\boldsymbol{h}(\boldsymbol{x}^*) = \boldsymbol{0}^\top .$$

□

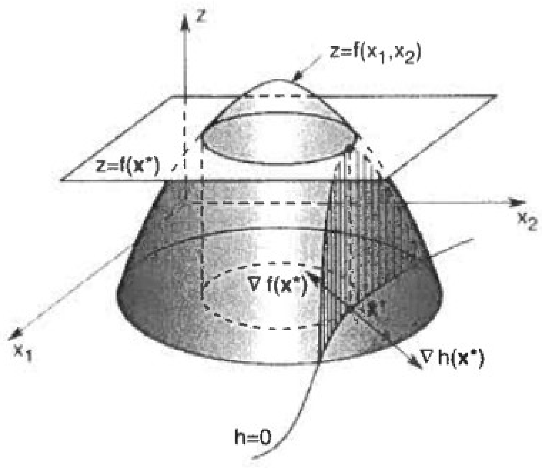


Figure 19.11 Lagrange's theorem for $n = 2$, $m = 1$.

Lagrange Condition

Given a fixed area of cardboard, we wish to construct a closed cardboard box with maximum volume. We can formulate and solve this problem using the Lagrange condition. Denote the dimensions of the box with maximum volume by x_1, x_2, and x_3, and let the given fixed area of cardboard be A. The problem can then be formulated as

$$\begin{array}{ll} \text{maximize} & x_1x_2x_3 \\ \text{subject to} & x_1x_2 + x_2x_3 + x_3x_1 = \dfrac{A}{2}. \end{array}$$

$$\begin{aligned} x_2x_3 - \lambda(x_2 + x_3) &= 0 \\ x_1x_3 - \lambda(x_1 + x_3) &= 0 \\ x_1x_2 - \lambda(x_1 + x_2) &= 0 \\ x_1x_2 + x_2x_3 + x_3x_1 &= \frac{A}{2}, \end{aligned}$$

Henrik Karstoft
Aarhus University, Department of Engineering

Lagrange Condition

Consider the problem of extremizing the objective function

$$f(\boldsymbol{x}) = x_1^2 + x_2^2$$

on the ellipse

$$\{[x_1, x_2]^\top : h(\boldsymbol{x}) = x_1^2 + 2x_2^2 - 1 = 0\}.$$

$$\nabla f(\boldsymbol{x}) = [2x_1, 2x_2]^\top,$$
$$\nabla h(\boldsymbol{x}) = [2x_1, 4x_2]^\top.$$

$$\begin{aligned} 2x_1 + 2\lambda x_1 &= 0, \\ 2x_2 + 4\lambda x_2 &= 0, \\ x_1^2 + 2x_2^2 &= 1. \end{aligned}$$

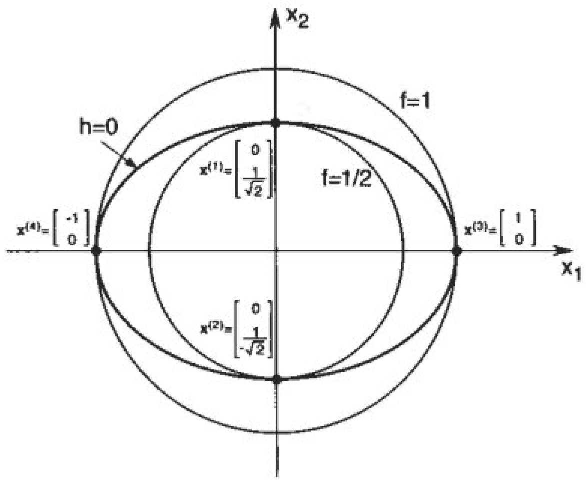


Figure 19.14 Graphical solution of the problem in Example 19.7.

Lagrange Condition

Consider the following problem:

$$\text{maximize} \quad \frac{\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{P}\boldsymbol{x}},$$

where $\boldsymbol{Q} = \boldsymbol{Q}^\top \geq 0$ and $\boldsymbol{P} = \boldsymbol{P}^\top > 0$. Note that if a point $\boldsymbol{x} = [x_1, \ldots, x_n]^\top$ is a solution to the problem, then so is any nonzero scalar multiple of it,

$$t\boldsymbol{x} = [tx_1, \ldots, tx_n]^\top, \quad t \neq 0.$$

The optimization problem becomes

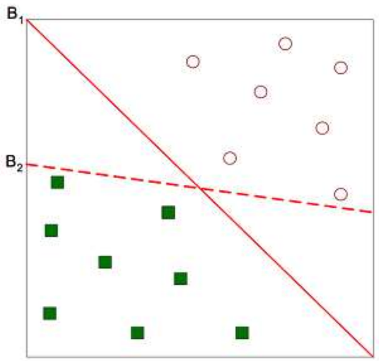
$$\begin{aligned} \text{maximize} \quad & \boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{x}^\top \boldsymbol{P}\boldsymbol{x} = 1. \end{aligned}$$

$$\boldsymbol{P}^{-1}\boldsymbol{Q}\boldsymbol{x} = \lambda \boldsymbol{x}.$$

$$\lambda^* = \boldsymbol{x}^{*\top} \boldsymbol{Q}\boldsymbol{x}^*.$$

Lagrange Condition, Support Vector Machines

Classifier (on feature space)



With decision boundary is the best?