Optimization Theory and Applications

Kun Zhu (zhukun@nuaa.edu.cn)

December 24, 2018

General Constrained Problems

- General problem with functional constraints:

$$\begin{array}{lll} \text{minimize} & f(\boldsymbol{x}) & \\ \text{subject to} & h_i(\boldsymbol{x}) = 0, & i = 1, \ldots, m \\ & g_j(\boldsymbol{x}) \leq 0, & j = 1, \ldots, p, \end{array}$$

  where $f : \mathbb{R}^n \to \mathbb{R}$, $h_i : \mathbb{R}^n \to \mathbb{R}$, $g_j : \mathbb{R}^n \to \mathbb{R}$, and $m < n$

- LP problem is an example of such a problem
- We will develop techniques for solving the above problems (similar to FONC, SONC, SOSC)

General Constrained Problems

- Example: Consider the problem

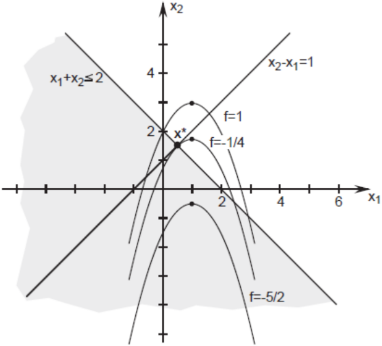
$$
\begin{array}{rl}
\text{minimize} & (x_1 - 1)^2 + x_2 - 2 \\
\text{subject to} & x_2 - x_1 = 1, \\
& x_1 + x_2 \leq 2.
\end{array}
$$

- The constraint (feasible) set is

$$S = \{\mathbf{x} \in \mathbb{R}^2 : x_2 - x_1 = 1, x_1 + x_2 \leq 2\}$$

General Constrained Problems

- We can solve this problem graphically



- In general, the graphical approach will not suffice. We need more powerful tools

Problems with Equality Constraints

- We now focus on problems with only equality constraints:

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & h_i(\boldsymbol{x}) = 0, \qquad i = 1, \ldots, m. \end{array}$$

- Write $\mathbf{h} = [h_1, \ldots, h_m]^T$, we can use vector notation:

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

where $\mathbf{h} : \mathbb{R}^n \to \mathbb{R}^m, m < n$

Problems with Equality Constraints

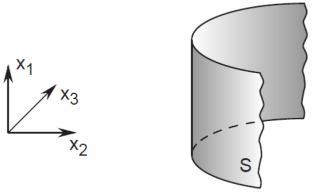
- We always assume that $f, \mathbf{h} \in \mathcal{C}^1$
- For simplicity, we first consider the case where $m = 1$. The constraint is $h(x) = 0$ (scalar)
- ***Definition*** ($m = 1$ case): A feasible point $\mathbf{x}^*$ is said to be *regular* if $\nabla h(\mathbf{x}^*) \neq 0$
- ”Regular” describes the smoothness of the curve or surface
- The feasible points define a surface. Define the surface as $S = \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) = 0\}$
- Geometrically, if all points in the constraint set S are regular, then the dimension of the surface S is $n - 1$

Problems with Equality Constraints

- Example: consider the constraint set

$$S = \{\mathbf{x} \in \mathbb{R}^3 : h_1(\mathbf{x}) = x_2 - x_3^2 = 0\}$$

- Here, $n = 3$ and $m = 1$
- We have $\nabla h_1(\mathbf{x}) = [0, 1, -2x_3]^T$, which is nonzero everywhere. Hence, any point in S is regular
- The dimension of S is $3 - 1 = 2$



$S=\{[x_1,x_2,x_3]^T : x_2-x_3^2=0\}$

Lagrange Conditions

- We now give a FONC type necessary condition for problems with equality constraints
- First consider the simple case where $m = 1$

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

where $f : \mathbb{R}^n \to \mathbb{R}, h : \mathbb{R}^n \to \mathbb{R}$

Lagrange Conditions

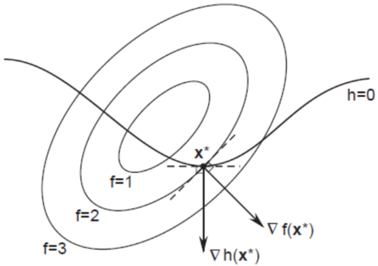
- ***Lagrange's Theorem*** ($m = 1$ case): suppose $\mathbf{x}^*$ is a local minimizer and is regular. Then there exists a scalar λ^* such that

$$\nabla f(\mathbf{x}^*) + \lambda^* \nabla h(\mathbf{x}^*) = 0$$

- In other words, $\nabla f(\mathbf{x}^*)$ and $\nabla h(\mathbf{x}^*)$ are parallel
- λ^* is called the *Lagrange multiplier*

Lagrange Conditions

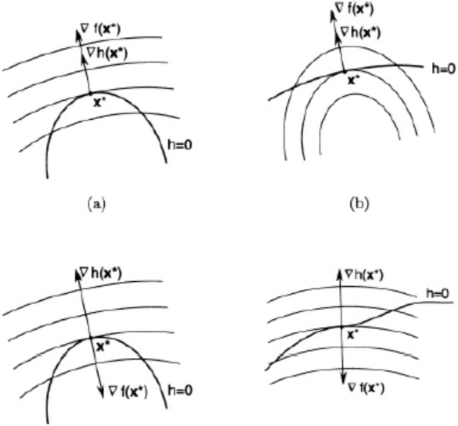
- Idea of proof of the theorem
  - Note that $\nabla f(\mathbf{x}^*)$ is orthogonal to the level set of f
  - Also, $\nabla h(\mathbf{x}^*)$ is orthogonal to the constraint set S
  - If $\nabla f(\mathbf{x}^*)$ and $\nabla h(\mathbf{x}^*)$ were not parallel, then we can move in a direction along S in the opposite direction to $\nabla f(\mathbf{x}^*)$, and the objective function decreases



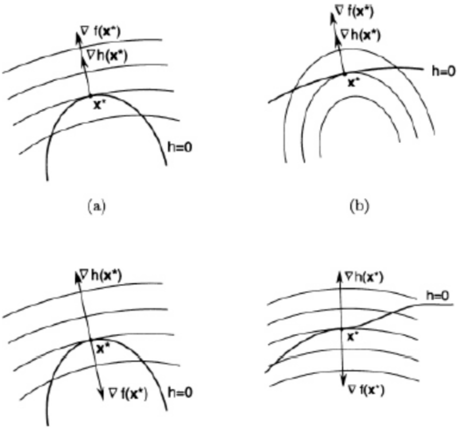
Lagrange Conditions

- Note that the Lagrange condition is only a necessary condition, not sufficient in general
- Since it is only a first order condition, both minimizers and maximizers satisfy it
- There may also be points that are neither minimizers nor maximizers that satisfy it

Lagrange Conditions



Lagrange Conditions



(a) maximizer (b) minimizer (c) minimizer (d) not an extreme point

Lagrange Conditions

- To apply the Lagrange theorem, it is convenient to define the *Lagrangian function*

$$l(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda h(\mathbf{x})$$

  Note that l is a function from $\mathbb{R}^{n+1}$ to $\mathbb{R}$

- Then, the condition in Lagrange's Theorem is equivalent to the FONC for l; i.e., we just take $\nabla l(\mathbf{x}^*, \lambda^*) = 0$

- To see this, note that

$$\nabla l(\mathbf{x}, \lambda) = \left[\begin{array}{c} \nabla_{\mathbf{x}} l(\mathbf{x}, \lambda) \\ \nabla_{\lambda} l(\mathbf{x}, \lambda) \end{array} \right]$$

Lagrange Conditions

- We have

$$\nabla_{\mathbf{x}} l(\mathbf{x}, \lambda) = \nabla f(\mathbf{x}) + \lambda \nabla h(\mathbf{x})$$
$$\nabla_{\lambda} l(\mathbf{x}, \lambda) = h(\mathbf{x})$$

- Therefore, the condition $\nabla l(\mathbf{x}^*, \lambda^*) = 0$ is equivalent to the two conditions

$$\nabla f(\mathbf{x}^*) + \lambda^* \nabla h(\mathbf{x}^*) = 0$$
$$h(\mathbf{x}^*) = 0$$

We call the above the ***Lagrange conditions***

Lagrange Conditions

- ***Example***: consider the optimization problem where

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

  and

$$h(\mathbf{x}) = x_1^2 + 2x_2^2 - 1$$

- The level sets are circles, and the constraint set is an eclipse
- We have $\nabla f(\mathbf{x}) = [2x_1, 2x_2]^T$, $\nabla h(\mathbf{x}) = [2x_1, 4x_2]^T$
- Note that all feasible points are regular

Lagrange Conditions

- The Lagrangian function
$$l(\mathbf{x}, \lambda) = x_1^2 + x_2^2 + \lambda(x_1^2 + 2x_2^2 - 1)$$
- To solve the problem, we first write down the Lagrange conditions ($\nabla_x l(\mathbf{x}, \lambda) = 0,\ \nabla_\lambda l(\mathbf{x}, \lambda) = 0$):
$$\begin{aligned} 2x_1 + 2\lambda x_1 &= 0 \\ 2x_2 + 4\lambda x_2 &= 0 \\ x_1^2 + 2x_2^2 &= 1 \end{aligned}$$
- From the first equation, we get $x_1 = 0$ or $\lambda = -1$
- For the case where $x_1 = 0$, the second and third equations imply that $\lambda = -1/2$ and $x_2 = \pm 1/\sqrt{2}$
- For the case where $\lambda = -1$, the second and third equations imply that $x_1 = \pm 1$ and $x_2 = 0$

Lagrange Conditions

- Accordingly, we find that there are four points that satisfy the Lagrange conditions

$$\boldsymbol{x}^{(1)} = \begin{bmatrix} 0 \\ 1/\sqrt{2} \end{bmatrix}, \quad \boldsymbol{x}^{(2)} = \begin{bmatrix} 0 \\ -1/\sqrt{2} \end{bmatrix}, \quad \boldsymbol{x}^{(3)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{x}^{(4)} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

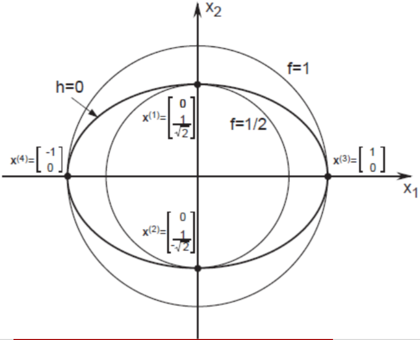
- All four points are regular

Lagrange Conditions

- Note that

$$f(\mathbf{x}^{*(1)}) = f(\mathbf{x}^{*(2)}) = 1/2 \quad \text{and} \quad f(\mathbf{x}^{*(3)}) = f(\mathbf{x}^{*(4)}) = 1$$

- Thus, if there are minimizers, then they are located at $\mathbf{x}^{*(1)}$ and $\mathbf{x}^{*(2)}$, and if there are maximizers, then they are located at $\mathbf{x}^{*(3)}$ and $\mathbf{x}^{*(4)}$



Lagrange Conditions

- ***Example***: Given a fixed area of cardboard (2 sq meter), we wish to construct a closed cardborad box with maximum volume. Find the dimensions of the (closed) box that has maximum volume
- ***Problem formulation***: denote the dimensions by x_1, x_2, and x_3. The problem is then

$$\begin{array}{ll} \text{maximize} & x_1x_2x_3 \\ \text{subject to} & x_1x_2 + x_2x_3 + x_3x_1 = 1 \end{array}$$

Lagrange Conditions

- Denote

$$\begin{aligned} f(\boldsymbol{x}) &= -x_1x_2x_3, \\ h(\boldsymbol{x}) &= x_1x_2 + x_2x_3 + x_3x_1 - 1 \end{aligned}$$

- We have

$$\nabla f(\boldsymbol{x}) = -\begin{bmatrix} x_2x_3 \\ x_1x_3 \\ x_1x_2 \end{bmatrix}, \qquad \nabla h(\boldsymbol{x}) = \begin{bmatrix} x_2 + x_3 \\ x_1 + x_3 \\ x_1 + x_2 \end{bmatrix}$$

- By the Lagrange conditions, the dimensions of the box with maximum volume satisfies

$$\begin{aligned} x_2x_3 - \lambda(x_2 + x_3) &= 0 \\ x_1x_3 - \lambda(x_1 + x_3) &= 0 \\ x_1x_2 - \lambda(x_1 + x_2) &= 0 \\ x_1x_2 + x_2x_3 + x_3x_1 - 1 &= 0, \end{aligned}$$

Lagrange Conditions

- x_1, x_2, x_3 are positive
- λ is nonzero (if $\lambda = 0$, then the constraints are violated)
- To solve the Lagrange conditions, first, multiply the first equation by x_1 and the second by x_2, and subtract one from the other
- We get $x_3\lambda(x_1 - x_2) = 0$, which implies that $x_1 = x_2$
- We similarly deduce that $x_2 = x_3$
- Hence, from the constraint equation, we deduce that

$$x_1^* = x_2^* = x_3^* = \frac{1}{\sqrt{3}}$$

Lagrange Conditions

- Note that regularity is stated as an assumption in Lagrange's theorem. This assumption plays an essential role
- ***Example***: consider the following problem:

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

where $f(x) = x$ and

$$h(x) = \begin{cases} x^2 & \text{if } x < 0 \\ 0 & \text{if } 0 \le x \le 1 \\ (x-1)^2 & \text{if } x > 1. \end{cases}$$

Lagrange Conditions

- The feasible set is evidently $[0, 1]$. And clearly $x^* = 0$ is a local minimizer
- However, $f'(x^*) = 1$ and $h'(x^*) = 0$
- Therefore, $x^* = 0$ is a minimizer that does not satisfy the necessary condition in Lagrange's theorem
- The reason is that x^* is not a regular point

Extension to General Case

- We now consider the general m case. The constraint is $\mathbf{h}(\mathbf{x}) = \mathbf{0}$
- ***Definition***: A feasible point $\mathbf{x}^*$ is said to be *regular* if $\nabla h_1(\mathbf{x}^*), \ldots, \nabla h_m(\mathbf{x}^*)$ are linear independent. That is $D\mathbf{h}(\mathbf{x}^*)$ is of full rank , where $D\mathbf{h}(\mathbf{x}^*)$ is the Jacobian matrix of $\mathbf{h}$ at $\mathbf{x}^*$

$$D\boldsymbol{h}(\boldsymbol{x}^*) = \begin{bmatrix} Dh_1(\boldsymbol{x}^*) \\ \vdots \\ Dh_m(\boldsymbol{x}^*) \end{bmatrix} = \begin{bmatrix} \nabla h_1(\boldsymbol{x}^*)^\top \\ \vdots \\ \nabla h_m(\boldsymbol{x}^*)^\top \end{bmatrix}$$

- The definition for $m = 1$ is a special case of the above
- Geometrically, if all points in the constraint set $\mathcal{S}$ are regular, then the dimension of $\mathcal{S}$ is $n - m$

Extension to General Case

- Example: For $n = 3$ and $m = 2$, consider the constraint $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, where

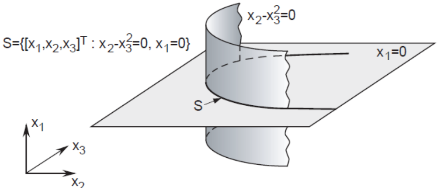
$$\mathbf{h}(\mathbf{x}) = \begin{bmatrix} x_2 - x_3^2 \\ x_1 \end{bmatrix}$$

- We have

$$D\mathbf{h}(\mathbf{x}) = \begin{bmatrix} 0 & 1 & -2x_3 \\ 1 & 0 & 0 \end{bmatrix}$$

  which has rank 2 everywhere. Hence, any point in $\mathcal{S}$ is regular

- The dimension of $\mathcal{S}$ is $3 - 2 = 1$



General Lagrange Theorem

- ***Lagrange Theorem***: Suppose $\mathbf{x}^*$ is a local minimizer and is regular. Then there exists $\lambda^* \in \mathbb{R}^m$ such that

$$Df(\mathbf{x}^*) + \lambda^{*T} D\mathbf{h}(\mathbf{x}^*) = \mathbf{0}^T$$

- $\lambda^* \in \mathbb{R}^m$ is called the Lagrange multiplier vector

General Lagrange Theorem

- As before, it is convenient to define the Lagrangian function

$$l(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda^T \mathbf{h}(\mathbf{x})$$

  Note that l is a function from $\mathbb{R}^{n+m}$ to $\mathbb{R}$

- Then, the Lagrange condition is equivalent to the FONC for l; i.e., we just take $\nabla l(\mathbf{x}^*, \lambda^*) = 0$

- The condition $\nabla l(\mathbf{x}^*, \lambda^*) = 0$ is equivalent to

$$\begin{aligned} Df(x^*) + \boldsymbol{\lambda}^{*T} Dh(x^*) &= 0^T \\ h(x^*) &= 0. \end{aligned}$$

General Lagrange Theorem

- Example: consider the problem

$$\begin{array}{ll} \text{minimize} & (1-x_1)^3 - (x_2+1)^2 + 3x_3^2 \\ \text{subject to} & x_1 = 0 \\ & x_2 = x_3^2. \end{array}$$

- Write

$$\begin{aligned} f(\boldsymbol{x}) &= (1-x_1)^3 - (x_2+1)^2 + 3x_3^2, \\ \boldsymbol{h}(\boldsymbol{x}) &= \begin{bmatrix} x_1 \\ x_2 - x_3^2 \end{bmatrix}. \end{aligned}$$

- We have

$$\begin{aligned} Df(\boldsymbol{x}) &= [-3(1-x_1)^2, -2(x_2+1), 6x_3], \\ D\boldsymbol{h}(\boldsymbol{x}) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2x_3 \end{bmatrix}. \end{aligned}$$

Note that all feasible points are regular

General Lagrange Theorem

- Introduce the vector $[\lambda_1, \lambda_2]^T$. The Lagrange conditions are

$$\begin{aligned} -3(1-x_1)^2 + \lambda_1 &= 0 \\ -2(x_2+1) + \lambda_2 &= 0 \\ 6x_3 - 2\lambda_2 x_3 &= 0 \\ x_1 &= 0 \\ x_2 - x_3^2 &= 0. \end{aligned}$$

- We immediately deduce that $x_1 = 0$, $\lambda_1 = 3$
- The third equation implies that either $x_3 = 0$ or $\lambda_2 = 3$
- Hence, there are three solutions:

$$\boldsymbol{x}^{*(1)} = \mathbf{0}, \text{ with } \boldsymbol{\lambda}^{*(1)} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$\boldsymbol{x}^{*(2)} = \begin{bmatrix} 0 \\ 1/2 \\ 1/\sqrt{2} \end{bmatrix}, \boldsymbol{x}^{*(3)} = \begin{bmatrix} 0 \\ 1/2 \\ -1/\sqrt{2} \end{bmatrix}, \text{ with } \boldsymbol{\lambda}^{*(2)} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}.$$

Tangent Space

- The tangent space at a point $\mathbf{x}^*$ on the surface $\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{h}(\mathbf{x}) = \mathbf{0}\}$ is the set

$$T(\mathbf{x}^*) = \{\mathbf{y} : D\mathbf{h}(\mathbf{x}^*)\mathbf{y} = \mathbf{0}\}$$

- Note that the tangent space $T(\mathbf{x}^*)$ is the nullspace of the matrix $D\mathbf{h}(\mathbf{x}^*)$; i.e.,

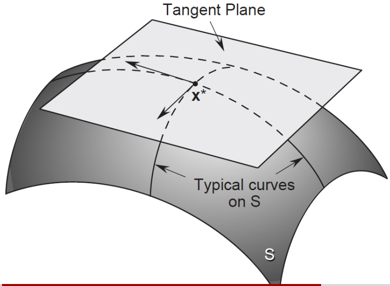
$$T(\mathbf{x}^*) = \mathcal{N}(D\mathbf{h}(\mathbf{x}^*))$$

- The tangent space is a subspace of $\mathbb{R}^n$ (plane passing through the origin)

- The dimension of $T(\mathbf{x}^*)$ is $n - m$ (assuming regularity of $\mathbf{x}^*$)

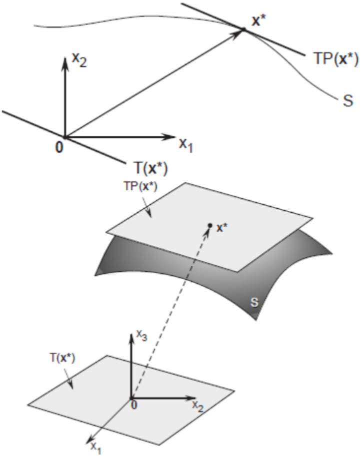
Tangent Space

- It is often convenient to picture the tangent space as a plane passing through the point $\mathbf{x}^*$
- For this, we define the *tangent plane* at $\mathbf{x}^*$ to be the set
$$TP(\mathbf{x}^*) = T(\mathbf{x}^*) + \mathbf{x}^*$$
- Geometric view: we shift $T(\mathbf{x}^*)$ so that it touches $\mathbf{x}^*$, then the resulting plane is tangent at $\mathbf{x}^*$. We call this plane the tangent plane



Tangent Space

- Geometric view: we shift $T(\mathbf{x}^*)$ so that it touches $\mathbf{x}^*$, the resulting plane is tangent at $\mathbf{x}^*$ (i.e., tangent plane)



Tangent Space

- Example: Let $\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^3 : h_1(\mathbf{x}) = x_1 = 0, h_2(\mathbf{x}) = x_1 - x_2 = 0\}$
- Then, $\mathcal{S}$ is the x_3-axis in $\mathbb{R}^3$
- We have

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

- ∇h_1 and ∇h_2 are linearly independent for any $\mathbf{x} \in \mathcal{S}$, all points of $\mathcal{S}$ are regular

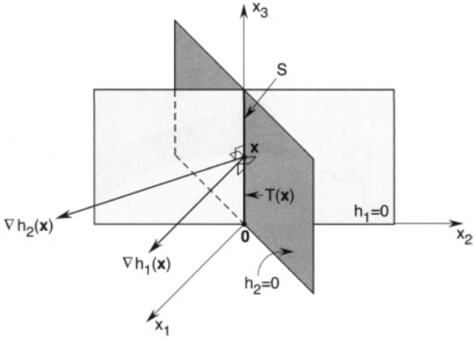
Tangent Space

- The tangent space at an arbitrary point of $\mathcal{S}$ is

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

Tangent Space

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



Second Order Conditions

- We now develop a SONC and SOSC for problems with equality constraints
- We assume that $f, \mathbf{h} \in \mathcal{C}^1$
- Recall: Lagrangian function

$$l(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^T \boldsymbol{h}(\boldsymbol{x}) = f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i h_i(\boldsymbol{x}).$$

- Given λ, the Hessian of $l(\mathbf{x}, \lambda)$ with respect to $\mathbf{x}$ is denoted

$$\boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \boldsymbol{F}(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i \boldsymbol{H}_i(\boldsymbol{x}),$$

where $\mathbf{F}$ is the Hessian of f, and $\mathbf{H}_i$ is the Hessian of h_i, $i = 1, \ldots, m$

Second Order Conditions

- Theorem: (SONC) Suppose $\mathbf{x}^*$ is a local minimizer and is regular. Then, there exists $\lambda^* \in \mathbb{R}^m$ such that the Lagrange conditions hold, and

$$\boldsymbol{y}^T \boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)\boldsymbol{y} \geq 0 \text{ for all } \boldsymbol{y} \in T(\boldsymbol{x}^*).$$

- Vectors $\mathbf{y} \in T(\mathbf{x}^*)$ play the role of “feasible directions”

Example of SONC

- Recall the example where

$$\begin{aligned} f(\boldsymbol{x}) &= (1-x_1)^3 - (x_2+1)^2 + 3x_3^2, \\ \boldsymbol{h}(\boldsymbol{x}) &= \begin{bmatrix} x_1 \\ x_2 - x_3^2 \end{bmatrix}. \end{aligned}$$

- We have

$$l(\boldsymbol{x}, \boldsymbol{\lambda}) = (1-x_1)^3 - (x_2+1)^2 + 3x_3^2 + \lambda_1 x_1 + \lambda_2 (x_2 - x_3^2).$$

- The Hessian (w.r.t. $\mathbf{x}$) of the Lagrangian is

$$\boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \begin{bmatrix} 6(1-x_1) & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 2(3-\lambda_2) \end{bmatrix}.$$

Example of SONC

- Recall that there are three solutions (all regular):

$$\boldsymbol{x}^{*(1)} = \mathbf{0}, \text{ with } \boldsymbol{\lambda}^{*(1)} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$\boldsymbol{x}^{*(2)} = \begin{bmatrix} 0 \\ 1/2 \\ 1/\sqrt{2} \end{bmatrix}, \boldsymbol{x}^{*(3)} = \begin{bmatrix} 0 \\ 1/2 \\ -1/\sqrt{2} \end{bmatrix}, \text{ with } \boldsymbol{\lambda}^{*(2)} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}.$$

- We will check if each of these three solutions satisfies the SONC

Example of SONC

- Recall that

$$Dh(\boldsymbol{x}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2x_3 \end{bmatrix}.$$

- Consider the first solution $\mathbf{x}^* = 0$ with $\lambda_2^{*(1)} = 2$. In this case,

$$\boldsymbol{L}(\boldsymbol{x}^{*(1)}, \boldsymbol{\lambda}^{*(1)}) = \begin{bmatrix} 6 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Example of SONC

- The tangent space is $T(\mathbf{x}^{*(1)}) = \mathcal{N}(D\mathbf{h}(\mathbf{x}^*))$, where

$$D\boldsymbol{h}(\boldsymbol{x}^{*(1)}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

- Hence, $T(\mathbf{x}^{*(1)}) = \mathbf{y} \in \mathbb{R}^3 : y_1 = y_2 = 0$, i.e., the x_3-axis
- Hence, the SONC holds at $\mathbf{x}^{*(1)} = 0$

Example of SONC

- Consider now the second solution $\mathbf{x}^{*(1)} = [0, 1/2, 1/\sqrt{2}]^T$ with $\lambda_2^* = [3, 3]^T$. In this case,

$$\boldsymbol{L}(\boldsymbol{x}^{*(2)}, \boldsymbol{\lambda}^{*(2)}) = \begin{bmatrix} 6 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

- The tangent space is $T(\mathbf{x}^{*(2)}) = \mathcal{N}(D\mathbf{h}(\mathbf{x}^{*2}))$, where

$$D\boldsymbol{h}(\boldsymbol{x}^{*(2)}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\sqrt{2} \end{bmatrix}.$$

Example of SONC

- Hence, $T(\mathbf{x}^{*(2)}) = \mathbf{y} : y_1 = 0, y_2 = y_3\sqrt{2}$
- In this case, we see that the SONC does not hold. Indeed, consider $\mathbf{y} = [0, \sqrt{2}, 1]^T$, we have

$$\boldsymbol{y}^T \boldsymbol{L}(\boldsymbol{x}^{*(2)}, \boldsymbol{\lambda}^{*(2)})\boldsymbol{y} = -4 \ngeq 0.$$

Second Order Conditions

- Theorem: (SOSC) Suppose $\mathbf{x}^*$ (feasible) and λ^* satisfy

  1. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}^T$; and
  2. $\boldsymbol{y}^T \boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)\boldsymbol{y} > 0$ for all nonzero $\boldsymbol{y} \in T(\boldsymbol{x}^*)$.

  Then $\mathbf{x}^*$ is a strict local minimizer