Problems with inequality constraints (Chap. 20)

- So far we have considered only problems with equality constraints: $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$.
- We now consider problems that have inequality constraints: $\boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}$, where $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^p$.
- As before, we give necessary conditions for problems with equality and inequality constraints.

Simple case: only inequality constraints

- Consider the problem

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

  where $\boldsymbol{g}(\boldsymbol{x}) = [g_1(\boldsymbol{x}), \ldots, g_p(\boldsymbol{x})]^T$.
- As usual, we assume $f, \boldsymbol{g} \in \mathcal{C}^1$.
- A point $\boldsymbol{x}$ is feasible iff $g_1(\boldsymbol{x}) \leq 0, \ldots, g_p(\boldsymbol{x}) \leq 0$.
- Definition: We say that the jth constraint $g_j \leq 0$ is *active* at $\boldsymbol{x}^*$ if $g_j(\boldsymbol{x}^*) = 0$. It is *inactive* if $g_j(\boldsymbol{x}^*) < 0$.

- Note that if a constraint is inactive at $\boldsymbol{x}^*$, then it is inactive at all points in some neighborhood of $\boldsymbol{x}^*$. Hence, locally around $\boldsymbol{x}^*$, the inactive constraints can be "ignored".
- Define $J(\boldsymbol{x}^*) = \{j : g_j(\boldsymbol{x}^*) = 0\}$, the set of indices of constraints that are active.
- Definition: A feasible point $\boldsymbol{x}^*$ is *regular* if the vectors $\nabla g_j(\boldsymbol{x}^*)$, $j \in J(\boldsymbol{x}^*)$, are linearly independent.

- Let $\boldsymbol{x}^*$ be a local minimizer of the original problem (with inequality constraint) and regular.
- Consider the optimization problem

$$\begin{array}{rl} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & g_j(\boldsymbol{x}) = 0,\ j \in J(\boldsymbol{x}^*) \end{array}$$

- Note that $\boldsymbol{x}^*$ is also a (regular) local minimizer for the above problem.
- Therefore, the Lagrange conditions hold at $\boldsymbol{x}^*$ for the above problem.

- Hence, by the Lagrange Theorem, there exists μ_j^*, $j \in J(\boldsymbol{x}^*)$, such that

$$Df(\boldsymbol{x}^*) + \sum_{j \in J(\boldsymbol{x}^*)} \mu_j^* Dg_j(\boldsymbol{x}^*) = \mathbf{0}^T.$$

- Let us define $\mu_j^* = 0$ for $j \notin J(\boldsymbol{x}^*)$ (i.e., all inactive j).

- Then, we can write the above condition as

$$Df(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^T,$$

  where $\boldsymbol{\mu}^* = [\mu_1^*, \ldots, \mu_p^*]^T$.

- Note that

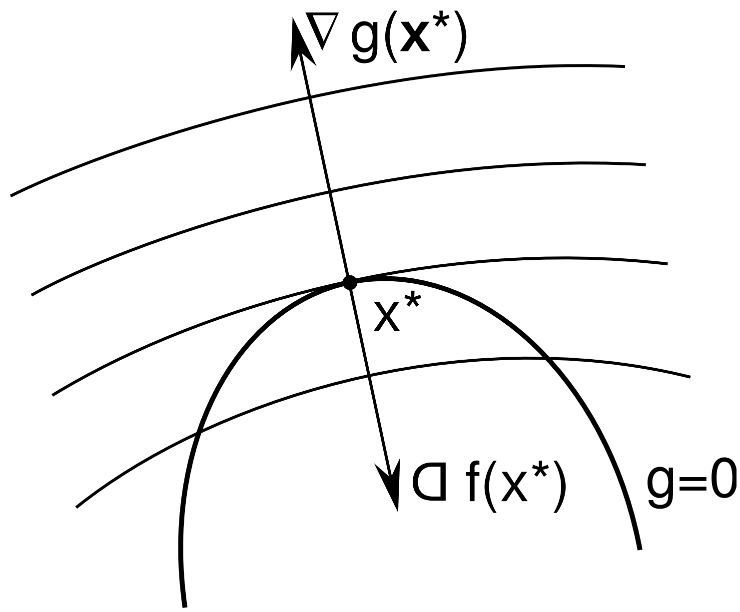
$$\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0,$$

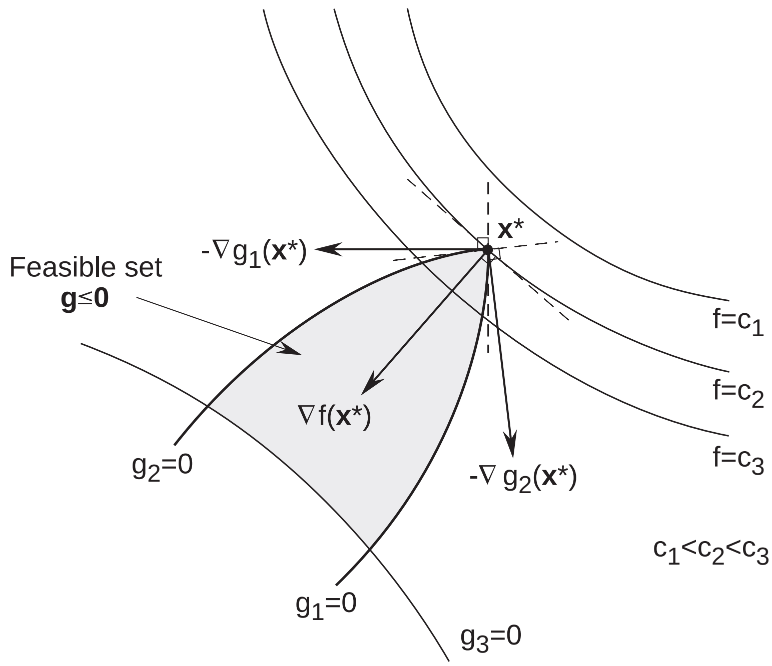
  because for each j, either $g_j(\boldsymbol{x}^*) = 0$ (active j) or $\mu_j^* = 0$ (inactive j). In other words, for all $j \notin J(\boldsymbol{x}^*)$ (inactive), we have $\mu_j^* = 0$.

- It turns out that we can say more about $\boldsymbol{\mu}^*$: every component of it is ≥ 0.

- To see this, we only need to concentrate on those $j \in J(\boldsymbol{x}^*)$, since the other μ_j^*'s are 0.

- It is easiest to illustrate the above fact using a picture.





- Note that the vectors $-\nabla g_j(\boldsymbol{x}^*)$, $j \in J(\boldsymbol{x}^*)$, point in the direction of the feasible region.
- Therefore, for $\boldsymbol{x}^*$ to be a local minimizer, $\nabla f(\boldsymbol{x}^*)$ must be a linear combination of the $-\nabla g_j(\boldsymbol{x}^*)$, $j \in J(\boldsymbol{x}^*)$, with nonnegative coefficients.
- This corresponds to:

$$\nabla f(\boldsymbol{x}^*) = -\sum_{j \in J(\boldsymbol{x}^*)} \mu_j^* \nabla g_j(\boldsymbol{x}^*),$$

  where $\mu_j^* \geq 0$.

Summary

- Consider the problem

$$\begin{array}{rl} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}. \end{array}$$

- Theorem (special case of Thm. 20.1): Suppose $\boldsymbol{x}^*$ is a local minimizer and is regular. Then, there exists $\boldsymbol{\mu}^* \in \mathbb{R}^p$ such that
  1. $\boldsymbol{\mu}^* \geq \boldsymbol{0}$
  2. $Df(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}^T$
  3. $\boldsymbol{\mu}^{*T} \boldsymbol{g}(\boldsymbol{x}^*) = 0$
  4. $\boldsymbol{g}(\boldsymbol{x}^*) \leq \boldsymbol{0}$.

- This theorem is called the *Karush-Kuhn-Tucker (KKT) Theorem*. The conditions are called KKT conditions (note that we usually include the constraints as part of the KKT conditions).
- The vector $\boldsymbol{\mu}^*$ is called the KKT multiplier vector.
- We have already seen the idea behind the proof.
- Note that for feasible $\boldsymbol{x}^*$ and $\boldsymbol{\mu}^*$,

$$\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0 \;\Leftrightarrow\; \mu_i^* g_i(\boldsymbol{x}^*) = 0 \text{ for all } i = 1, \ldots, p$$

- Actually, there is a more general version of the theorem, where we have both equality and inequality constraints (see later).

Example: (20.3)

- Consider the problem

$$\begin{array}{ll} \text{minimize} & x_1^2 + x_2^2 + x_1x_2 - 3x_1 \\ \text{subject to} & x_1, x_2 \geq 0, \end{array}$$

- The KKT conditions for this problem are
  1. $\boldsymbol{\mu} = [\mu_1, \mu_2]^T \geq \mathbf{0}$;
  2. $Df(\boldsymbol{x}) - \boldsymbol{\mu}^T = \mathbf{0}^T$;
  3. $\boldsymbol{\mu}^T\boldsymbol{x} = 0$.
  4. $\boldsymbol{x} \geq \mathbf{0}$.

  (Note: slightly different from book.)

- We have $Df(\boldsymbol{x}) = [2x_1 + x_2 - 3,\; x_1 + 2x_2]$. Note that all feasible points are regular here.
- Hence, we have

$$\begin{aligned} 2x_1 + x_2 - \mu_1 &= 3 \\ x_1 + 2x_2 - \mu_2 &= 0 \\ \mu_1x_1 + \mu_2x_2 &= 0 \\ \mu_1, \mu_2, x_1, x_2 &\geq 0. \end{aligned}$$

- We have four variables, three equations, and inequality constraints on each variable.
- To find a solution for $\boldsymbol{x}^*$, $\boldsymbol{\mu}^*$, we first notice that it is impossible for $\boldsymbol{x}^* = \mathbf{0}$.

- Focus on the third equation. First try $x_1^* = 0$.
- By the first equation, we must have $x_2^* > 0$. Thus, $\mu_2^* = 0$.
- Solving the equations we obtain

$$x_2^* = 0, \qquad \mu_1^* = -3,$$

which is not valid.

- Next, we try $x_2^* = 0$, which then implies $\mu_1^* = 0$.
- Solving, we obtain

$$x_1^* = \frac{3}{2}, \qquad \mu_2^* = \frac{3}{2},$$

which is evidently a valid solution to the KKT conditions.

- What about the case $x_1^* > 0$ and $x_2^* > 0$? Not possible (why?).
- Note that to solve conditions that have inequalities, we have to "try" solutions that are at the boundary (active constraints). [Recall LP feasibility problem.]

Maximization and/or $\boldsymbol{g}(\boldsymbol{x}) \geq \mathbf{0}$

- We can easily modify the KKT conditions to problems with maximization or inequality constraints of the form $\boldsymbol{g}(\boldsymbol{x}) \geq \mathbf{0}$.
- In the case of maximization, either we change the sign f, or we can change the sign of $\boldsymbol{\mu}^*$.
- Similarly, in the case of constraints of the form $\boldsymbol{g}(\boldsymbol{x}) \geq \mathbf{0}$, either we change the sign $\boldsymbol{g}$, or we can change the sign of $\boldsymbol{\mu}^*$.

- Specifically, consider the problem

$$\begin{array}{ll} \text{maximize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}. \end{array}$$

- The KKT conditions for the above problem are
  1. $\boldsymbol{\mu}^* \leq \mathbf{0}$
  2. $Df(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^T$
  3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$
  4. $\boldsymbol{g}(\boldsymbol{x}^*) \leq \mathbf{0}$

  (The only difference is the sign of $\boldsymbol{\mu}^*$.)

- Similarly, for the problem

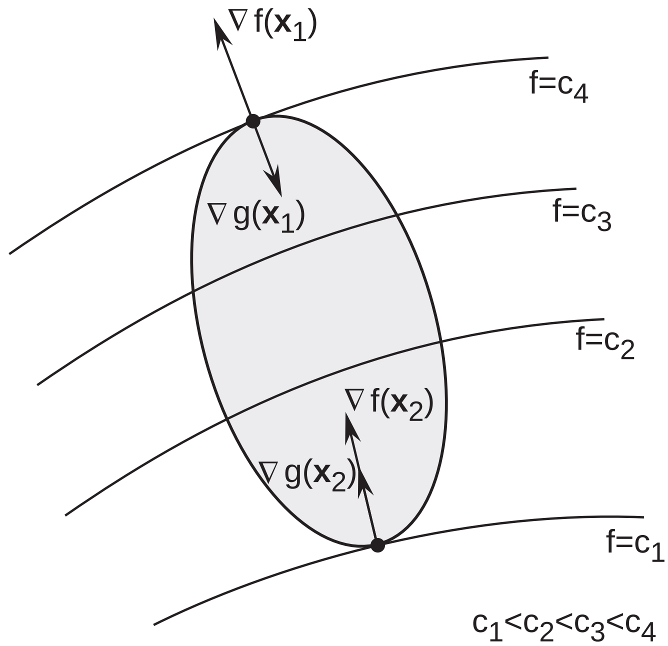
$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0}, \end{array}$$

  the KKT conditions are

  1. $\boldsymbol{\mu}^* \leq \boldsymbol{0}$
  2. $Df(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}^T$
  3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$
  4. $\boldsymbol{g}(\boldsymbol{x}^*) \geq \boldsymbol{0}$

- If we have both maximization and $\boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0}$, then the KKT conditions are the same as the original (standard) case [except for the constraint].

For the case $\boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0}$:



Example:

- Consider the general problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \geq \boldsymbol{0}. \end{array}$$

- The KKT conditions are

$$\boldsymbol{\mu} \leq \boldsymbol{0}$$

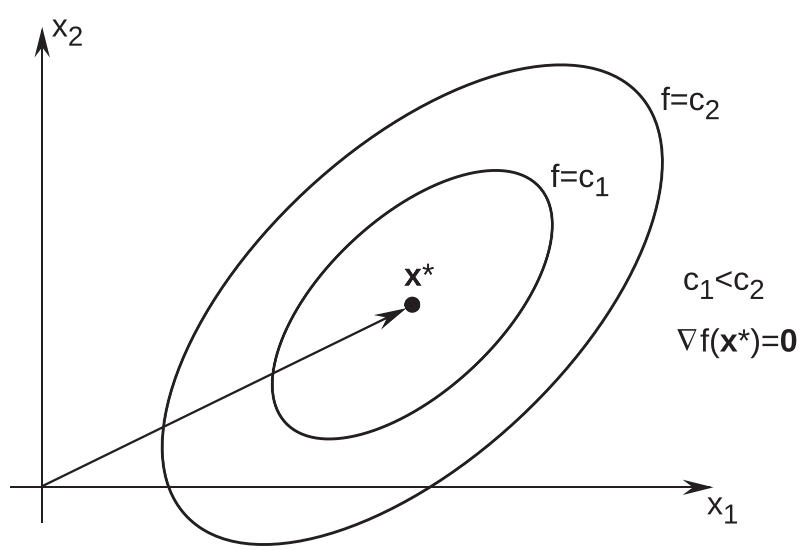
$$\begin{aligned} \nabla f(\boldsymbol{x}) + \boldsymbol{\mu} &= \mathbf{0} \\ \boldsymbol{\mu}^T \boldsymbol{x} &= 0 \\ \boldsymbol{x} &\geq \mathbf{0}. \end{aligned}$$

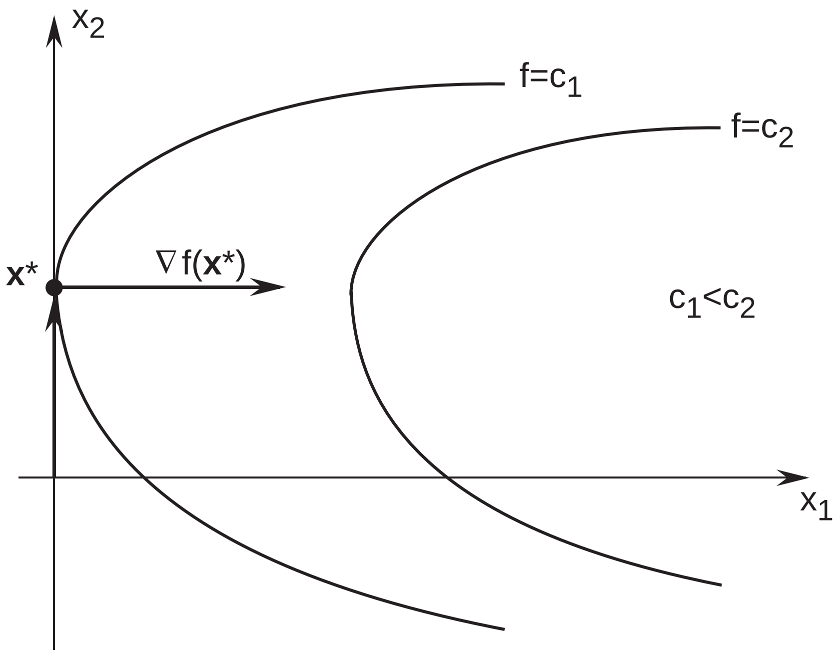
Note that we have used the appropriately modified KKT conditions (with $\boldsymbol{\mu} \leq \mathbf{0}$). Also note that all feasible points are regular here.

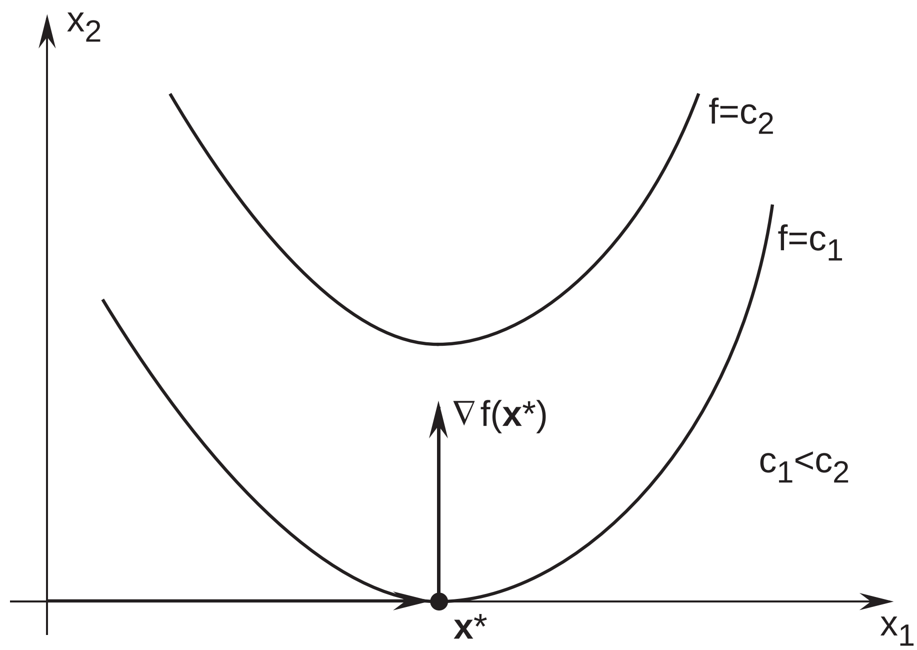
- From KKT conditions, we easily deduce that

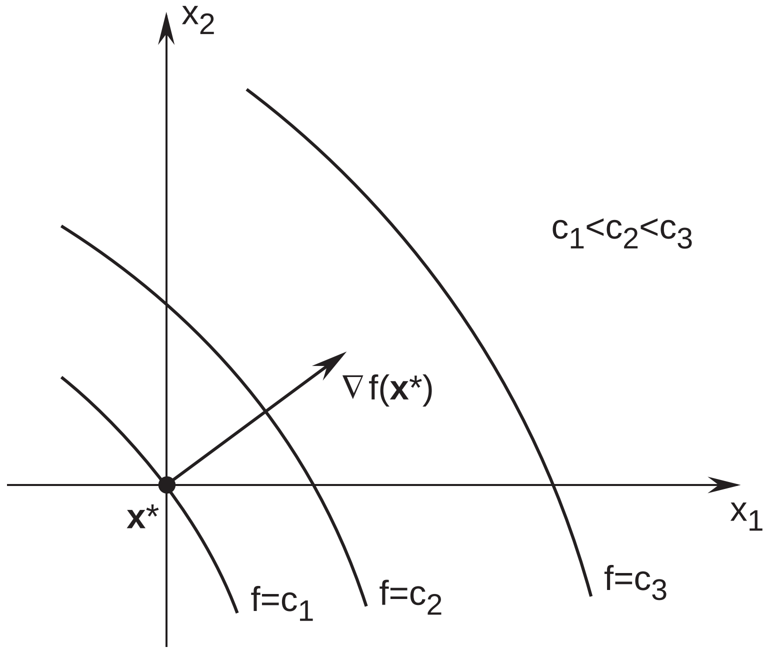
$$\begin{aligned} \nabla f(\boldsymbol{x}) &\geq \mathbf{0} \\ \boldsymbol{x}^T \nabla f(\boldsymbol{x}) &= 0 \\ \boldsymbol{x} &\geq \mathbf{0}. \end{aligned}$$

- Compare the above conditions to the FONC for set constraints (involving *feasible directions*).
- We can illustrate some possible points in $\mathbb{R}^2$ that satisfy the above conditions. Which ones are local minimizers?




x2
f=c1
f=c2
x*
∇f(x*)
c1<c2
x1


x2
f=c2
f=c1
∇f(x*)
c1<c2
x*
x1



Problems with equality and inequality constraints (§20.1)

- Consider the optimization problem

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}, \\ & \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}, \end{aligned}$$

  where $f : \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{h} : \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$, and $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^p$.

- Our goal is to derive necessary conditions for the above general problem.

- Definition: A feasible point $\boldsymbol{x}^*$ is *regular* if the vectors

$$\nabla h_i(\boldsymbol{x}^*),\ i = 1, \ldots, m,\ \nabla g_j(\boldsymbol{x}^*), j \in J(\boldsymbol{x}^*),$$

  are linearly independent.

- By convention we consider every equality constraint $h_i = 0$ to be *active*.

- Hence, regularity means the gradients of all active constraint functions are linearly independent.

- Theorem (general KKT, Thm. 20.1): Suppose $\boldsymbol{x}^*$ is a local minimizer and is regular. Then, there exists $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ and $\boldsymbol{\mu}^* \in \mathbb{R}^p$ such that

  1. $\boldsymbol{\mu}^* \geq \mathbf{0}$

2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}^T$
3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$
4. $\boldsymbol{h}(\boldsymbol{x}^*) = \boldsymbol{0}$
5. $\boldsymbol{g}(\boldsymbol{x}^*) \leq \boldsymbol{0}$.

- The difference between the above KKT conditions and the previous one (with no equality constraints) is that we need to incorporate the Lagrange multiplier vector $\boldsymbol{\lambda}^*$.

- The idea behind the proof of the (general) KKT Theorem is the same as what we have seen for the special case with no equality constraints.

- Basically, the proof involves applying the Lagrange Theorem to the associated problem with only equality constraints involving active constraints at $\boldsymbol{x}^*$:

$$\begin{array}{rl} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \\ & g_j(\boldsymbol{x}) = 0,\ j \in J(\boldsymbol{x}^*), \end{array}$$

and, as before, we have $\boldsymbol{\mu}^* \geq \boldsymbol{0}$ and $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$.

Example: Savings in bank

- Bank interest paid monthly at rate $r > 0$ (compound).

- We wish to deposit some money into the bank every month for n months, such that the total is D dollars.

- Goal: maximize the total amount of money accumulated at the end of n months.

- Let x_i be amount deposited in beginning of ith month;

- Optimization problem:

$$\begin{array}{rl} \text{maximize} & (1+r)^n x_1 + (1+r)^{n-1}x_2 + \cdots + (1+r)x_n \\ \text{subject to} & x_1 + \cdots + x_n = D \\ & x_1, \ldots, x_n \geq 0 \end{array}$$

- Write

$$\begin{aligned} f(\boldsymbol{x}) &= -\big((1+r)^n x_1 + (1+r)^{n-1}x_2 + \cdots + (1+r)x_n\big) \\ h(\boldsymbol{x}) &= x_1 + \cdots + x_n - D \\ \boldsymbol{g}(\boldsymbol{x}) &= -[x_1, \ldots, x_n]^T = -\boldsymbol{x}. \end{aligned}$$

- We have

$$\begin{aligned} Df(\boldsymbol{x}) &= -[(1+r)^n, (1+r)^{n-1}, \dots, (1+r)] \\ Dh(\boldsymbol{x}) &= [1, 1, \dots, 1] \\ D\boldsymbol{g}(\boldsymbol{x}) &= -\boldsymbol{I}_n. \end{aligned}$$

- The KKT conditions are

$$\begin{aligned} \mu_1, \dots, \mu_n &\geq 0 \\ -(1+r)^{n-i+1} + \lambda - \mu_i &= 0, \ i = 1, \dots, n \\ \mu_1 x_1 + \cdots + \mu_n x_n &= 0 \\ x_1 + \cdots + x_n &= D \\ x_1, \dots, x_n &\geq 0. \end{aligned}$$

- Suppose that $x_1^* > 0$. Then, $\mu_1^* = 0$, and so we have

$$\begin{aligned} \lambda^* &= (1+r)^n, \\ \mu_i^* &= (1+r)^n - (1+r)^{n-i+1} > 0, \ i = 2, \dots, n, \\ x_1^* &= D, \ x_i^* = 0, \ i = 2, \dots, n. \end{aligned}$$

- The previous solution corresponds to depositing D dollars in the first month.
- Are there any other solutions?
- Suppose $x_i^* > 0$, $i \neq 1$ (hence $\mu_i^* = 0$). We then conclude that

$$\begin{aligned} \lambda^* &= (1+r)^{n-i+1}, \\ \mu_{i-1}^* &= (1+r)^{n-i+1} - (1+r)^{n-i+2} < 0, \end{aligned}$$

  which is clearly not valid.
- Hence, there are no other solutions.

- The KKT Theorem provides a necessary condition (not sufficient in general).
- Therefore, based only on the KKT Theorem, we cannot tell whether the solution obtained is in fact a minimizer. (However, since there is only one solution, it must be the global minimizer if indeed one exists).
- The above example problem is an LP problem. It turns out that our solution is a BFS for which the associated RCCs are all ≥ 0. Hence, by our results from LP, we can prove that our solution is a global minimizer (Exercise: try this).

Second order conditions (§20.2)

- As before, we can develop second-order conditions for local minimizers.
- We assume that $f, \boldsymbol{h}, \boldsymbol{g} \in \mathcal{C}^2$.
- We can now derive a SONC for problems with equality and inequality constraints.
- The idea is to apply the SONC for problems with only equality constraints to the problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \\ & g_j(\boldsymbol{x}) = 0, \ j \in J(\boldsymbol{x}^*). \end{array}$$

- Let us define (similar to before)

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

  where $\boldsymbol{F}$ is the Hessian of f, $\boldsymbol{H}_i$ is the Hessian of h_i, $i = 1, \ldots, m$, and $\boldsymbol{G}_i$ is the Hessian of g_i, $i = 1, \ldots, p$.

- We will also need the tangent space for the surface defined by the active constraints:

$$T(\boldsymbol{x}^*) = \\ \{\boldsymbol{y} \in \mathbb{R}^n : D\boldsymbol{h}(\boldsymbol{x}^*)\boldsymbol{y} = \boldsymbol{0}, \ Dg_j(\boldsymbol{x}^*)\boldsymbol{y} = 0, \ j \in J(\boldsymbol{x}^*)\}.$$

- Theorem (20.2): (SONC) Suppose $\boldsymbol{x}^*$ is a local minimizer and is regular. Then, there exists $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$ such that the KKT conditions hold, and

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

- As before, we often use the terminology "$\boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \geq 0$ on $T(\boldsymbol{x}^*)$" to refer to the above condition.

Example:

- Consider the previous example (bank deposits).
- The only point satisfying the KKT conditions is $\boldsymbol{x}^* = [D, 0, \ldots, 0]^T$, with $\lambda^* = (1 + r)^n$, and $\mu_i^* = (1 + r)^n - (1 + r)^{n-i+1}$, $i = 1, \ldots, n$.
- For this problem,

$$\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{O}, \ \boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{O}, \ \boldsymbol{G}_i(\boldsymbol{x}) = \boldsymbol{O}.$$

  Thus,

$$\boldsymbol{L}(\boldsymbol{x}^*, \lambda^*, \boldsymbol{\mu}^*) = \boldsymbol{O}.$$

- Thus, the SONC holds (trivially).

- Even though we don't need to compute $T(\boldsymbol{x}^*)$ in this case, let us do it anyway, for completeness.

- We have $J(\boldsymbol{x}^*) = \{2, 3, \ldots, n\}$.

- Hence,

$$\begin{aligned} T(\boldsymbol{x}^*) &= \{\boldsymbol{y} : [1, \ldots, 1]\boldsymbol{y} = 0,\ y_i = 0,\ i = 2, \ldots, n\} \\ &= \{\boldsymbol{0}\}. \end{aligned}$$

- Therefore, regardless of $\boldsymbol{L}(\boldsymbol{x}^*, \lambda^*, \boldsymbol{\mu}^*)$, the SONC holds in this case.

Example

- Consider the problem

$$\begin{array}{ll} \text{minimize} & -\dfrac{1}{2}\left((x_1 - 1)^2 + x_2^2\right) \\ \text{subject to} & x_1, x_2 \geq 0. \end{array}$$

- First write

$$\begin{aligned} f(\boldsymbol{x}) &= -\frac{1}{2}\left((x_1 - 1)^2 + x_2^2\right), \\ \boldsymbol{g}(\boldsymbol{x}) &= -[x_1, x_2]^T = -\boldsymbol{x}. \end{aligned}$$

- Hence,

$$Df(\boldsymbol{x}) = -[x_1 - 1, x_2], \qquad D\boldsymbol{g}(\boldsymbol{x}) = -\boldsymbol{I}_2.$$

- The KKT conditions are

$$\begin{aligned} x_1 - 1 + \mu_1 &= 0 \\ x_2 + \mu_2 &= 0 \\ \mu_1 x_1 + \mu_2 x_2 &= 0 \\ \mu_1, \mu_2, x_1, x_2 &\geq 0. \end{aligned}$$

- There are two solutions to the above conditions.

- The first solution is:

$$\mu_1^* = \mu_2^* = 0, \qquad x_1^* = 1, x_2^* = 0.$$

- We have $\boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\mu}^*) = -\boldsymbol{I}_2$ and

$$\begin{aligned} T(\boldsymbol{x}^*) &= \{\boldsymbol{y} : [0,1]\boldsymbol{y} = 0\} \\ &= \{\boldsymbol{y} : y_2 = 0\} \\ &= x_1\text{-axis.} \end{aligned}$$

- Hence, the SONC does not hold for this solution (hence it is not a local minimizer).

- The second solution is:

$$\mu_1^* = 1, \mu_2^* = 0, \qquad x_1^* = x_2^* = 0.$$

- We have $\boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\mu}^*) = -\boldsymbol{I}_2$ (again) and

$$T(\boldsymbol{x}^*) = \left\{\boldsymbol{y} : \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{y} = \boldsymbol{0}\right\} = \{\boldsymbol{0}\}.$$

- Hence, the SONC holds for this solution.
- Of course, we still cannot say for sure whether or not this solution is a local minimizer. We need a sufficient condition to be sure.

- Recall SONC, which requires checking whether $\boldsymbol{L}(\boldsymbol{x}^*, \lambda^*, \boldsymbol{\mu}^*) \geq 0$ on $T(\boldsymbol{x}^*)$.
- Are there situations where the SONC holds but the solution is not a local minimizer? Yes!
- What we need is a sufficient condition (SOSC).

- Define $\tilde{J}(\boldsymbol{x}^*, \boldsymbol{\mu}^*) = \{i : g_i(\boldsymbol{x}^*) = 0, \mu_i^* > 0\}$, i.e., those active constraints that are "nondegenerate".
- We also need to define another subspace:

$$\begin{aligned} &\tilde{T}(\boldsymbol{x}^*, \boldsymbol{\mu}^*) = \\ &\qquad \{\boldsymbol{y} : D\boldsymbol{h}(\boldsymbol{x}^*)\boldsymbol{y} = \boldsymbol{0}, Dg_i(\boldsymbol{x}^*)\boldsymbol{y} = 0, i \in \tilde{J}(\boldsymbol{x}^*, \boldsymbol{\mu}^*)\} \end{aligned}$$

- Note that $\tilde{J}(\boldsymbol{x}^*, \boldsymbol{\mu}^*) \subset J(\boldsymbol{x}^*)$.
- Hence, $T(\boldsymbol{x}^*) \subset \tilde{T}(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$.

- Theorem (20.3): (SOSC) Suppose $\boldsymbol{x}^*$ (feasible), $\boldsymbol{\lambda}^*$, and $\boldsymbol{\mu}^* \geq \boldsymbol{0}$ satisfy

  1. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}^T$,
     $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$; and

2. $\boldsymbol{y}^T \boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)\boldsymbol{y} > 0$
   for all nonzero $\boldsymbol{y} \in \tilde{T}(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$.

Then, $\boldsymbol{x}^*$ is a strict local minimizer.

- As usual, we often say "$\boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) > 0$ on $\tilde{T}(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$" to refer to condition 2 above.

- Note that condition 2 is more stringent that the SONC in two ways:
  1. We have > 0 rather than ≥ 0.
  2. The condition > 0 must hold on the larger subspace $\tilde{T}(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$ (rather than just $T(\boldsymbol{x}^*)$).

- Therefore, in situations where $T(\boldsymbol{x}^*) = \{\boldsymbol{0}\}$ (in which case the SONC holds trivially), but the matrix $\boldsymbol{L} = \boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ is not positive definite, and $\tilde{T}(\boldsymbol{x}^*, \boldsymbol{\mu}^*) \neq \{\boldsymbol{0}\}$, then the SOSC may not hold.

- Or, if $\boldsymbol{y}^T\boldsymbol{L}\boldsymbol{y} = 0$ for all $\boldsymbol{y} \in T(\boldsymbol{x}^*)$ (in which case the SONC holds), but $\boldsymbol{y}^T\boldsymbol{L}\boldsymbol{y} \leq 0$ for some nonzero $\boldsymbol{y} \in \tilde{T}(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$, then the SOSC does not hold.

- Note that if $\tilde{T}(\boldsymbol{x}^*, \boldsymbol{\mu}^*) = \{\boldsymbol{0}\}$, the SOSC holds automatically, regardless of $\boldsymbol{L}$.

Example

- Recall the previous example:

$$\begin{array}{ll} \text{minimize} & -\dfrac{1}{2}\left((x_1-1)^2 + x_2^2\right) \\ \text{subject to} & x_1, x_2 \geq 0. \end{array}$$

- We have $f(\boldsymbol{x}) = -\frac{1}{2}\left((x_1-1)^2 + x_2^2\right)$ and $\boldsymbol{g}(\boldsymbol{x}) = -[x_1, x_2]^T$.

- The following solution satisfies the FONC and SONC:

$$\mu_1^* = 1, \mu_2^* = 0, \qquad x_1^* = x_2^* = 0.$$

- Recall: $\boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\mu}^*) = -\boldsymbol{I}_2$.

- We have

$$\begin{aligned} \tilde{T}(\boldsymbol{x}^*, \boldsymbol{\mu}^*) &= \{\boldsymbol{y} : [1, 0]\boldsymbol{y} = \boldsymbol{0}\} \\ &= \{\boldsymbol{y} : y_1 = 0\} \\ &= x_2\text{-axis}. \end{aligned}$$

- So, $\boldsymbol{y} = [0, 1]^T \in \tilde{T}(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$ and $\boldsymbol{y}^T\boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\mu}^*)\boldsymbol{y} = -1 \not\geq 0$.

- Hence, the SOSC does *not* hold in this case.

