Nonlinear Optimization

Nonlinear Optimization Models

$$
\begin{array}{ll}
\min & f(x_1, x_2, \ldots, x_n) \\
\text{s.t.} & (x_1, x_2, \ldots, x_n) \in X
\end{array}
$$

Nonlinear Optimization

- Nonlinear Programming
  - **Types of Nonlinear Programs (NLP)**
  - **Convexity and Convex Programs**
  - **NLP Solutions**
- Unconstrained Optimization
  - **Principles of Unconstrained Optimization**
  - **Search Methods**
- Constrained Optimization Theory
  - **The KKT Conditions**
  - **The Lagrange Multiplier (Sensitivity Analysis)**
- Linearly Constrained Convex Optimization (LCCP)
  - **Duality and optimality conditions revisited**
  - **Solution concepts for Quadratic Programs (QP) and LCP**
- Classification of NLP Algorithms and Solution Methods

Nonlinear Programming

$$
\begin{aligned}
\min \quad & f(x_1, x_2, \dots, x_n) \\
\text{s.t.} \quad & \\
& c_1(x_1, x_2, \dots, x_n) \geq (=)\ 0 \\
& c_2(x_1, x_2, \dots, x_n) \geq (=)0 \\
& \quad \dots \qquad \dots \\
& c_m(x_1, x_2, \dots, x_n) \geq (=)\ 0
\end{aligned}
$$

Gradient Vector and Hessian Matrix

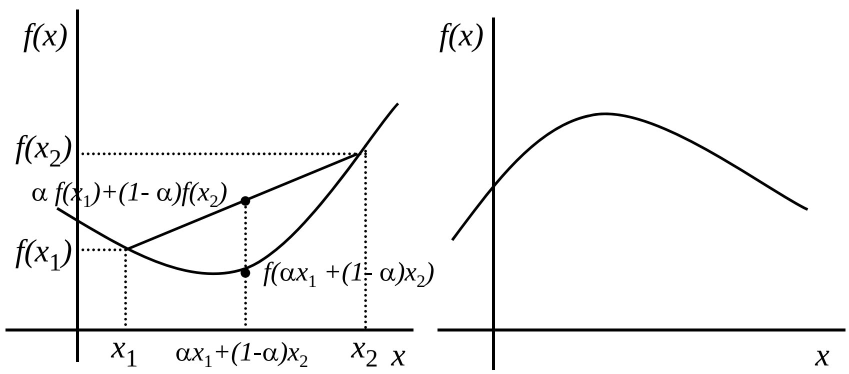
- The gradient vector of f at x

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \cdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

- The Hessian Matrix of f at x

$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}$$

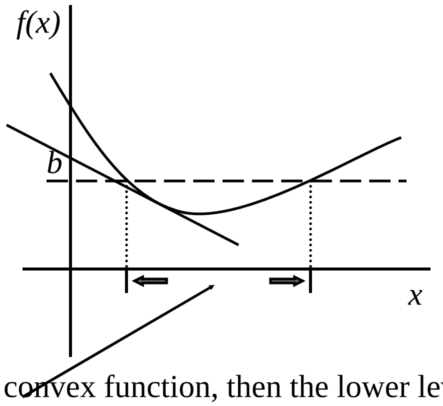
Convex and Concave Functions



f(x) is a convex function if and only if for any given two points x_1 and x_2 in the function domain and for any constant $0 \le \alpha \le 1$

$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2)$

Properties of Convex Function



If $f(x)$ is a convex function, then the lower level set $\{x: f(x) \leq b\}$ is a convex set for any constant b.

The graph of a convex function lies above its tangent planes.
The Hessian matrix of a convex function is positive semi-definite.

Convex Quadratic Function

$f(x)=x^TQx+c^Tx$ is a convex function if and only if Q is positive semi-definite.

$f(x)=x^TQx+c^Tx$ is a strictly convex function if and only if Q is positive definite.

If Q is positive definite, Q^{-1} exists.

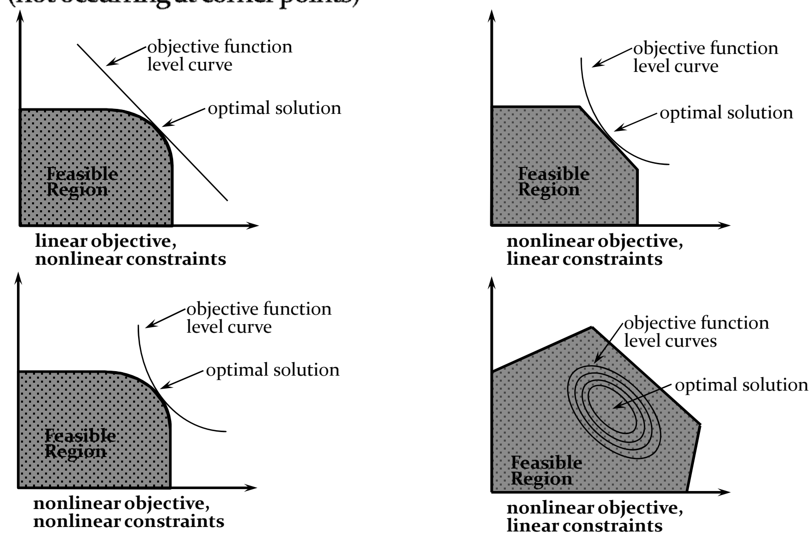
Convex Sets

- A set is <u>convex</u> if every line segment connecting any two points in the set is contained entirely within the set
  - Ex - polyhedron
  - Ex - ball
- An <u>extreme point</u> of a convex set is any point that is not on any line segment connecting any other two points of the set
- The intersection of convex sets is a convex set

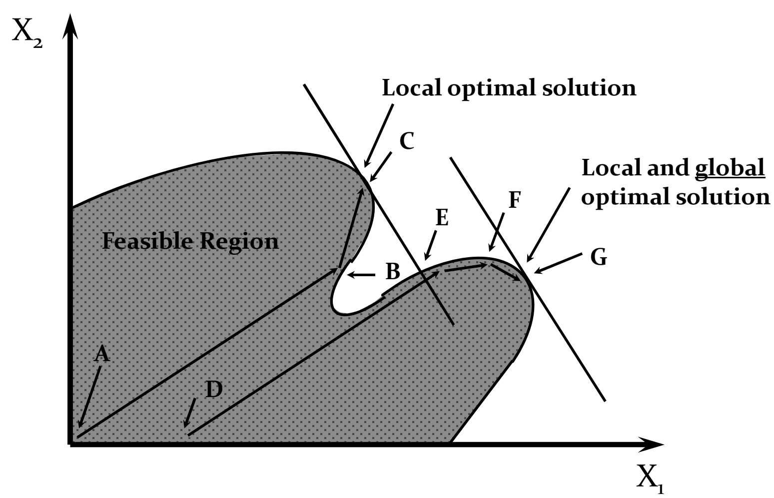
Why do we care so much about convexity?

- Because it guarantees that a local optimum is a global optimum
- This is significant because all of our basic optimization algorithms search for local optima
  - Those that try harder to find global optima generally just run underlying algorithms several times starting at different solutions

Possible Optimal Solutions to Convex NLPs
(not occurring at corner points)



Local vs. Global Optimal Solutions for Nonconvex NLPs



KKT Conditions:

- KKT conditions may not lead directly to a very efficient algorithm for solving NLPs. However, they do have a number of benefits:
  - They give insight into what optimal solutions to NLPs look like
  - They provide a way to set up and solve small problems
  - They provide a method to check solutions to large problems
  - The Lagrangian values can be seen as shadow prices of the constraints

Unconstrained Optimization

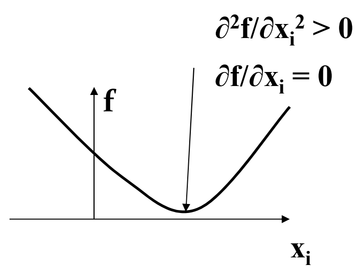
$$\min \quad f(x) = (x_1, x_2, \ldots, x_n)$$

Common Assumptions

- We generally assume *f(x)* is continuous and differentiable over the feasible region.
  - If it is not, we can still apply solution techniques but they often become a bit more complicated (e.g. have to examine at discontinuities)

Geometries of KKT: Unconstrained

- Problem:
  - **Minimize f(x), where x is a vector that could have any values, positive or negative**
- First Order Necessary Condition (min or max):
  - **$\nabla f(x) = 0$ ($\partial f/\partial x_i = 0$ for all i) is the first order necessary condition for optimization**
- Second Order Necessary Condition:
  - **$\nabla^2 f(x)$ is positive semidefinite (PSD)**
    - **[$x \cdot \nabla^2 f(x) \cdot x \geq 0$ for all x]**
- Second Order Sufficient Condition (Given FONC satisfied)
  - **$\nabla^2 f(x)$ is positive definite (PD)**
    - **[$x \cdot \nabla^2 f(x) \cdot x > 0$ for all x]**



Constrained optimization

- We will study how to characterize local optima for
  - multi-dimensional optimization problems
  - with more complex constraints
- We will start by considering problems with *only* equality constraints
  - We will also assume that the objective and constraint functions are continuous and differentiable

Constrained optimization: equality constraints

- A general *equality constrained* multi-dimensional NLP is:

$$
\begin{aligned}
&\max\ f(x) = f(x_1, \ldots, x_n) \\
&\text{subject to} \\
&\qquad g_1(x_1, \ldots, x_n) = b_1 \\
&\qquad g_2(x_1, \ldots, x_n) = b_2 \\
&\qquad \vdots \\
&\qquad g_m(x_1, \ldots, x_n) = b_m
\end{aligned}
$$

Constrained optimization: equality constraints

- The Lagrangian approach is to associate a *Lagrange multiplier* λ_i with the i^{th} constraint
- We then form the *Lagrangian* by adding weighted constraint violations to the objective function:

$$
\begin{aligned}
&L(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) = \\
&\qquad f(x_1, \ldots, x_n) + \sum_{i=1}^{m} \lambda_i \left(b_i - g_i(x_1, \ldots, x_n)\right)
\end{aligned}
$$

or $L(x, \lambda) = f(x) + \lambda' \left(b - g(x)\right)$

Constrained optimization: equality constraints

- Now consider the *stationary points* of the Lagrangian:

$$\frac{\partial L(x,\lambda)}{\partial x_j} = \frac{\partial f(x)}{\partial x_j} - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(x)}{\partial x_j} = 0 \qquad j = 1,\ldots,n$$
$$\frac{\partial L(x,\lambda)}{\partial \lambda_i} = b_i - g_i(x) = 0 \qquad i = 1,\ldots,m$$

  - The 2nd set of conditions says that x needs to satisfy the equality constraints!
  - The 1st set of conditions generalizes the unconstrained stationary point condition!

Constrained optimization: equality constraints

- Let (x^*,λ^*) maximize the Lagrangian
- Then it should be a stationary point of L
  - $g(x^*)=b$, i.e., x^* is a feasible solution to the original optimization problem
  - Furthermore, for all feasible x and all λ

$$L(x^*,\lambda^*) \geq L(x,\lambda)$$
$$f(x^*) + \lambda^{*\prime}\left(b - g(x^*)\right) \geq f(x) + \lambda'\left(b - g(x)\right)$$
$$f(x^*) \geq f(x)$$

Constrained optimization: equality constraints

- Conclusion: we can find the optimal solution to the *constrained* problem by considering all stationary points of the *unconstrained* Lagrangian problem
  - i.e., by finding all solutions to

$$\frac{\partial L(x,\lambda)}{\partial x_j} = \frac{\partial f(x)}{\partial x_j} - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(x)}{\partial x_j} = 0 \qquad j = 1,\dots,n$$

$$\frac{\partial L(x,\lambda)}{\partial \lambda_i} = b_i - g_i(x) = 0 \qquad i = 1,\dots,m$$

Constrained optimization: equality constraints

- As a byproduct, we get the interesting observation that

$$L(x^*,\lambda^*) = f(x^*) + \lambda^{*\prime}\left(b - g(x^*)\right) = f(x^*)$$

- We will use this later when interpreting the values of the multipliers λ^*

Constrained optimization: equality constraints

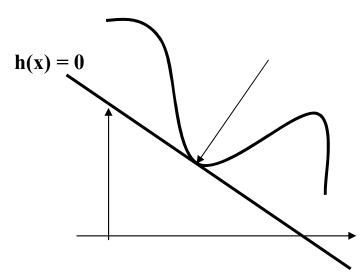
- Note: if
  - the objective function *f* is concave
  - all constraint functions g_i are linear
- Then any stationary point of *L* is an optimal solution to the constrained optimization problem!!
  - this result also holds when *f* is convex

Geometries of KKT: Equality Constrained (one constraint)

- Problem:
  - **Minimize f(x), where x is a vector**
  - **Subject to: h(x) = 0**
- First Order Necessary Condition for minimum (or for maximum):
  - $\nabla f(x) - \lambda \nabla h(x) = 0$

  for some λ free (λ is a scalar)

Two surfaces must be tangent



h(x) = 0 and -h(x) = 0 are the same; there is no sign restriction on λ

Geometries of KKT: Equality Constrained (one constraint)

- First Order Necessary Condition:
  - $\nabla f(x) - \lambda \nabla h(x) = 0$ for some λ
- Lagrangian:
  - $L(x, \lambda) = f(x) - \lambda [h(x)]$,
  - Minimize $L(x, \lambda)$ over x and Maximize $L(x, \lambda)$ over λ. Use principles of unconstrained optimization
  - $\nabla L(x, \lambda) = 0$:
  - $\nabla_x L(x, \lambda) = \nabla f(x) - \lambda \nabla h(x) = 0$
  - $\nabla_\lambda L(x, \lambda) = h(x) = 0$

Constrained optimization: equality constraints

- An example:

$$\begin{aligned} &\max \quad -x_1^2 - x_2^2 \\ &\text{s.t. } x_1 + x_2 = 1 \end{aligned}$$

Constrained optimization: equality constraints

- Then $$L(x_1, x_2, \lambda) = -x_1^2 - x_2^2 + \lambda(1 - x_1 - x_2)$$ $$= -x_1^2 - x_2^2 + \lambda - \lambda x_1 - \lambda x_2$$
- First order conditions:

$$\begin{aligned} -2x_1 - \lambda &= 0 \\ -2x_2 - \lambda &= 0 \\ 1 - x_1 - x_2 &= 0 \end{aligned} \quad \text{or} \quad \begin{aligned} x_1 &= -\tfrac{1}{2}\lambda \\ x_2 &= -\tfrac{1}{2}\lambda \\ x_1 + x_2 &= 1 \end{aligned} \quad \text{or} \quad \begin{aligned} x_1 &= \tfrac{1}{2} \\ x_2 &= \tfrac{1}{2} \\ \lambda &= -1 \end{aligned}$$

Constrained optimization: sensitivity analysis

- Recall that we found earlier that

$$L(x^*, \lambda^*) = f(x^*) + \lambda^{*\prime}\left(b - g(x^*)\right) = f(x^*)$$

- What happens to the optimal solution value if the right-hand side of constraint *i* is changed by a *small* amount, say Δb_i
  - It changes by *approximately* $\lambda_i^* \Delta b_i$
  - Compare this to sensitivity analysis in LP
  - λ_i^* is the *shadow price* of constraint *i*

Constrained optimization: sensitivity analysis

- LINGO:
- For a maximization problem, LINGO reports the values of λ_i at the *local* optimum found in the DUAL PRICE column
- For a minimization problem, LINGO reports the values of $-\lambda_i$ at the *local* optimum found in the DUAL PRICE column

Example 5: Advertising

- Q&H company advertises on soap operas and football games
  - Each soap opera ad costs $50,000
  - Each football game ad costs $100,000
- Q&H wants exactly 40 million men and 60 million women to see its ads
- How many ads should Q&H purchase in each category?

Example 5 (contd.): Advertising

- Decision variables:
  - S = number of soap opera ads
  - F = number of football game ads
- If S soap opera ads are bought, they will be seen by

$$5 \cdot 10^6 \sqrt{S} \text{ men and } 20 \cdot 10^6 \sqrt{S} \text{ women}$$

- If F football game ads are bought, they will be seen by

$$17 \cdot 10^6 \sqrt{F} \text{ men and } 7 \cdot 10^6 \sqrt{F} \text{ women}$$

Example 5 (contd.): Advertising

- Model:

$$\begin{aligned}
&\text{minimize } 50S + 100F \\
&\text{subject to} \\
&\qquad 5\sqrt{S} + 17\sqrt{F} = 40 \\
&\qquad 20\sqrt{S} + 7\sqrt{F} = 60 \\
&\qquad\qquad S, F \geq 0
\end{aligned}$$

Example 5 (contd.): Advertising

- LINGO:
  min=50*S+100*F;
  5*S^.5+17*F^.5=40;
  20*S^.5+7*F^.5=60;

Example 5 (contd.): Advertising

- Solution:

Local optimal solution found at iteration: 18
Objective value: 563.0744

Variable	Value	Reduced Cost
S	5.886590	0.000000
F	2.687450	0.000000

Row	Slack or Surplus	Dual Price
1	563.0744	-1.000000
2	0.000000	-15.93120
3	0.000000	-8.148348

Example 5 (contd.): Advertising

- *Interpretation:*
- How does the optimal cost change if we require that 41 million men see the ads?
- We have a minimization problem, so the Lagrange multiplier of the first constraint is approximately 15.931
  - Thus the optimal cost will increase by *approximately* \$15,931 to approximately \$579,005
  - (N.B., reoptimization of the modified problem yields an optimal cost of \$579,462)

Constrained optimization: Inequality Constraints

- We will still assume that the objective and constraint functions are continuous and differentiable
- We will assume all constraints are "$\leq$" constraints
- We will also look at problems with both equality and inequality constraints

Constrained optimization: inequality constraints

- A general *inequality constrained* multi-dimensional NLP is:

$$
\begin{aligned}
&\max \ f(x) = f(x_1, \ldots, x_n) \\
&\text{subject to} \\
&\qquad g_1(x_1, \ldots, x_n) \le b_1 \\
&\qquad g_2(x_1, \ldots, x_n) \le b_2 \\
&\qquad \vdots \\
&\qquad g_m(x_1, \ldots, x_n) \le b_m
\end{aligned}
$$

Constrained optimization: inequality constraints

- In the case of inequality constraints, we also associate a *multiplier* λ_i with the i^{th} constraint
- As in the case of equality constraints, these multipliers can be interpreted as *shadow prices*

Constrained optimization: inequality constraints

- Without derivation or proof, we will look at a set of *necessary* conditions, called Karush-Kuhn-Tucker- or KKT-conditions, for a given point, say $\bar{X}$, to be an optimal solution to the NLP
- These are valid when a certain condition ("constraint qualification") is verified.
- The latter will be assumed for now.

Constrained optimization: inequality constraints

- By necessity, an optimal point should satisfy the KKT-conditions.
- However, not all points that satisfy the KKT-conditions are optimal!

- The characterization holds under certain conditions on the constraints
  - The so-called "constraint qualification conditions"
  - in most cases these are satisfied
    - for example: if all constraints are linear

Constrained optimization: KKT conditions

- If $\bar{x}$ is an optimal solution to the NLP (in max-form), it must be feasible, and:
  - there must exist a vector of multipliers $\bar{\lambda}$ satisfying

$$
\begin{aligned}
\frac{\partial f(\bar{x})}{\partial x_j} - \sum_{i=1}^{m} \bar{\lambda}_i \frac{\partial g_i(\bar{x})}{\partial x_j} &= 0 \quad j = 1, \ldots, \mathrm{n} \\
\bar{\lambda}_i \left(b_i - g_i(\bar{x})\right) &= 0 \quad i = 1, \ldots, m \\
g_i(\bar{x}) &\le b_i \quad i = 1, \ldots, m \\
\bar{\lambda}_i &\ge 0 \quad i = 1, \ldots, m
\end{aligned}
$$

Constrained optimization: KKT conditions

- The second set of KKT conditions is

$$\bar{\lambda}_i \cdot \left(b_i - g_i(\bar{x})\right) = 0 \qquad i = 1, \ldots, m$$

- This is comparable to the *complementary slackness* conditions from LP!

$$\text{if } \bar{\lambda}_i > 0 \text{ then } g_i(\bar{x}) = b_i$$

$$\text{if } g_i(\bar{x}) < b_i \text{ then } \bar{\lambda}_i = 0$$

Constrained optimization: KKT conditions

- This can be interpreted as follows:
- Additional units of the *resource* b_i only have value if the available units are used fully in the optimal solution

- Finally, note that increasing b_i enlarges the feasible region, and therefore increases the objective value
  - Therefore, $\lambda_i \geq 0$ for all i

Constrained optimization: KKT conditions

- Derive similar sets of KKT-conditions for
  - A minimization problem
  - A problem having ≥-constraints
  - A problem having a mixture of constraints (≤, =, ≥)

Constrained optimization: sufficient conditions

- If
  - f is a *concave* function
  - $g_1,\ldots,g_m$ are *convex* functions

  then any solution x^* satisfying the KKT conditions is an optimal solution to the NLP
- A similar result can be formulated for *minimization* problems

Constrained optimization: inequality constraints

- An example:

$$\begin{aligned} \text{max} \quad & -x_1^2 - x_2^2 \\ \text{s.t.} \quad & -x_1 - x_2 \leq -1 \end{aligned}$$

Constrained optimization: inequality constraints

- The KKT conditions are:

$$
\begin{aligned}
-2x_1 - \lambda &= 0 \\
-2x_2 - \lambda &= 0 \\
\lambda \cdot (-1 + x_1 + x_2) &= 0 \\
-x_1 - x_2 &\le -1 \\
\lambda &\ge 0
\end{aligned}
$$

with solution

$$
\begin{aligned}
x_1 &= \tfrac{1}{2} \\
x_2 &= \tfrac{1}{2} \\
\lambda &= 1
\end{aligned}
$$

Constrained optimization: inequality constraints

- With multiple inequality constraints:

$$
\begin{aligned}
\max\quad & -x_1^2 - x_2^2 \\
\text{s.t.}\quad & -2x_1 - x_2 \le -1 \\
& -x_1 - 2x_2 \le -1
\end{aligned}
$$

Constrained optimization: inequality constraints

- The KKT conditions are:

$$
\begin{aligned}
-2x_1 + 2\lambda_1 + \lambda_2 &= 0 \\
-2x_2 + \lambda_1 + 2\lambda_2 &= 0 \\
\lambda_1 \cdot \left(-1 + 2x_1 + x_2\right) &= 0 \\
\lambda_2 \cdot \left(-1 + x_1 + 2x_2\right) &= 0 \\
-2x_1 - x_2 &\le -1 \\
-x_1 - 2x_2 &\le -1 \\
\lambda_1, \lambda_2 &\ge 0
\end{aligned}
$$

with solution

$$
\begin{aligned}
x_1 &= \tfrac{1}{3} \\
x_2 &= \tfrac{1}{3} \\
\lambda_1 &= \tfrac{2}{3} \\
\lambda_2 &= \tfrac{2}{3}
\end{aligned}
$$

Constrained optimization: inequality constraints

- Another example:

$$
\begin{aligned}
\text{max} \quad & -x_1^2 - x_2^2 \\
\text{s.t.} \quad & -5x_1 - x_2 \le -5 \\
& -3x_1 - 2x_2 \le -6
\end{aligned}
$$

Constrained optimization: inequality constraints

- The KKT conditions are:

$$
\begin{aligned}
2x_1 + 5\lambda_1 + 3\lambda_2 &= 0 \\
2x_2 + \lambda_1 + 2\lambda_2 &= 0 \\
\lambda_1 \cdot (-5 + 5x_1 + x_2) &= 0 \\
\lambda_2 \cdot (-6 + 3x_1 + 2x_2) &= 0 \\
-5x_1 - x_2 &\le -5 \\
-3x_1 - 2x_2 &\le -6 \\
\lambda_1, \lambda_2 &\ge 0
\end{aligned}
$$

with solution

$$
\begin{aligned}
x_1 &= 1\tfrac{5}{13} \\
x_2 &= \tfrac{12}{13} \\
\lambda_1 &= 0 \\
\lambda_2 &= \tfrac{12}{13}
\end{aligned}
$$

A Word on Constraint Qualification:

- It has to be satisfied before we can apply KKT theorem
- It comes in several flavors
- We only focus on the following:
  - The gradients of the constraint functions, including those corresponding to non-negativity, have to be linearly independent
- When the constraints are all linear, the constraint qualification is satisfied.

A Word on Constraint Qualification:

- An example:

$$\begin{aligned} \max \quad & f(x_1, x_2) = x_1 \\ s.t. \quad & x_2 - (1 - x_1)^3 \le 0 \qquad (1) \\ & x_1 \ge 0 \\ & x_2 \ge 0 \end{aligned}$$

A Word on Constraint Qualification:

If $x_1 > 1$ then inequality (1) above implies $x_2 < 0$.
Therefore x_1 must be ≤ 1 for a feasible solution.
Since $(x_1, x_2) = (1,0)$ is feasible and $f(x_1, x_2) = x_1$,
the maximum is achieved at $(1,0)$.

Geometry of KKT

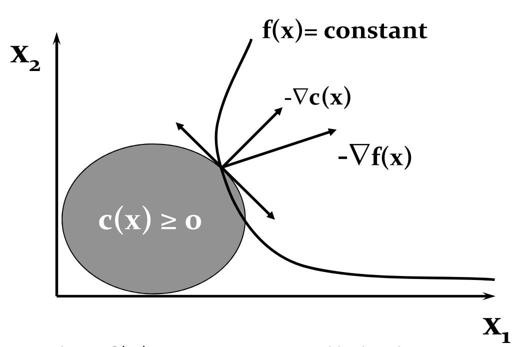
Geometries of KKT: Equality Constrained (multiple constraints)

- Problem:
  - Minimize f(x), where x is a vector
  - Such that: $h_i(x) = 0$ for i = 1,2,…,m
- KKT Conditions (necessary nonditions when constraint qualification holds):
  - $\nabla f(x) = \Sigma_{i=1}^{n} \lambda_i \nabla h_i(x)$
  - $h_i(x) = 0$ for i = 1,2,…,m
- Such a point (x, λ) is called a KKT point, and λ is called the Dual Vector or the Lagrange Multipliers. Furthermore, these conditions are sufficient if f(x) is convex and $h_i(x)$, i = 1,2,…,m, are linear.

Geometry of KKT: Inequality Constrained (one constraint)

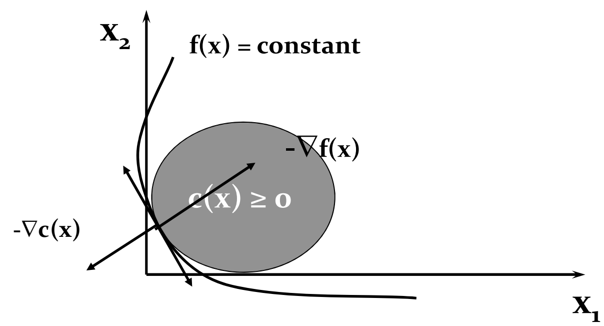
- Problem:
  - Minimize f(x), where x is a vector
  - Subject to: $c(x) \geq 0$.
- Assume feasible set and set of points preferred to any point are all convex sets.
  - (i.e. convex program)
- <u>First Order Necessary Condition</u>:
  - $\nabla f(x) = \lambda \nabla c(x)$ for some $\lambda \geq 0$ (λ is a scalar)
  - If constraint is binding [$c(x) = 0$], then $\lambda \geq 0$
  - If constraint is none-binding [$c(x) > 0$], then $\nabla f(x) = 0$ or $\lambda = 0$

Geometries of KKT: Inequality Constrained (one constraint)



If $-\nabla c(x)$ and $-\nabla f(x)$ are not parallel, there are feasible points with less f(x).

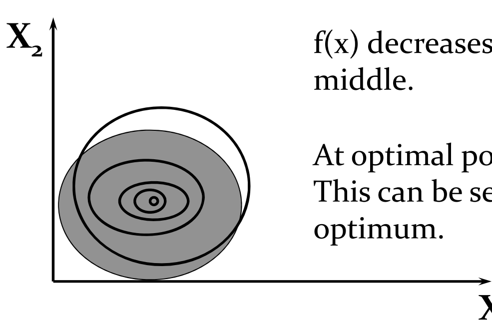
Geometries of KKT: Inequality Constrained (one constraint)



If -∇c(x) and -∇f(x) are parallel but in opposite direction, there are feasible points with less f(x).

Geometries of KKT: Inequality Constrained (one constraint)

- First Order Necessary Condition:
  - ∇f(x) = 0 if constraint is not binding [c(x) > 0]

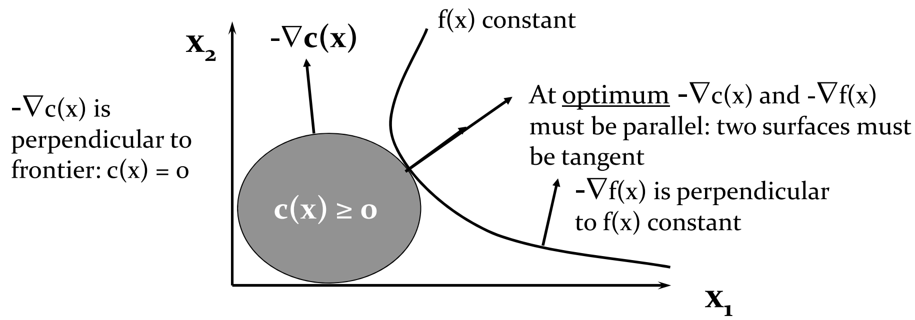


f(x) decreases toward sink at the middle.

At optimal point, ∇f(x) = 0
This can be sees as an unconstrained optimum.

Geometries of KKT: Inequality Constrained (one constraint)

- First Order Necessary Condition:
  - $\nabla f(x) = \lambda \nabla c(x)$ for some $\lambda \geq 0$ (λ is a scalar)
  - If constraint is binding [$c(x) = 0$] then $\lambda > 0$



Optimality Conditions (Multiple Constraints)

- Problem:
  - Minimize $f(x)$, where x is a vector
  - Such that: $c_i(x) \geq 0$ for $i = 1,2,\ldots,m$
- <u>KKT Conditions (necessary nonditions when constraint qualification holds):</u>
  - $\nabla f(x) = \Sigma_{i=1}^{m} \lambda_i \nabla c_i(x)$
  - $c_i(x) \geq 0$ for $i = 1,2,\ldots,m$
  - $\lambda_i \geq 0$ for $i = 1,2,\ldots,m$
  - $\lambda_i [c_i(x)] = 0$, for $i=1,2,\ldots,m$
- Furthermore, these conditions are sufficient if we are dealing with a convex programming problem

Example: KKT Conditions

min $x_1^2 + x_2^2$

s.t.

$$-0.25 \cdot (x_1 - 2)^2 - (x_2 - 2)^2 \geq -1$$

KKT Conditions

$$\begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} - \lambda \cdot \begin{pmatrix} 0.5(2 - x_1) \\ 2(2 - x_2) \end{pmatrix} = 0$$

$$-0.25 \cdot (x_1 - 2)^2 - (x_2 - 2)^2 \geq -1$$

$$\lambda \geq 0$$

$$\lambda[1 - 0.25 \cdot (x_1 - 2)^2 - (x_2 - 2)^2] = 0$$

Example: Computation of the KKT Condition

$$\begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} - \lambda \cdot \begin{pmatrix} 0.5(2 - x_1) \\ 2(2 - x_2) \end{pmatrix} = 0$$

$$x_1 = \frac{2\lambda}{4 + \lambda}; \qquad x_2 = \frac{2\lambda}{1 + \lambda}$$

- If $\lambda = 0$, then $x_1 = 0$ and $x_2 = 0$, and thus the constraint would not hold with equality. Therefore, λ must be positive.
- Plugging the two values of $x_1(\lambda)$ and $x_2(\lambda)$ into the constraint with equality gives us $\lambda = 1.8$.
- We can then solve for $x_1 = .61$ and $x_2 = 1.28$.

KKT Conditions: Final Notes

- KKT conditions may not lead directly to a very efficient algorithm for solving NLPs. However, they do have a number of benefits:
  - They give insight into what optimal solutions to NLPs look like
  - They provide a way to set up and solve small problems
  - They provide a method to check solutions to large problems
  - The Lagrange multipliers can be seen as shadow prices of the constraints