Constrained Nonlinear Optimization

Moritz Diehl & Sébastien Gros

Outline

Outline

Algebraic Characterization of **Unconstrained** Local Optima

Consider the unconstrained problem: $\min_{\mathbf{w}} \quad \Phi(\mathbf{w})$

1st-Order **Necessary** Condition of Optimality (FONC)

$$\mathbf{w}^* \text{ local optimum} \quad \Rightarrow \quad \nabla\Phi(\mathbf{w}^*) = \mathbf{0}, \; \mathbf{w}^* \text{ stationary point}$$

2nd-Order **Sufficient** Conditions of Optimality (SOSC)

NLP:

$$\nabla\Phi(\mathbf{w}^*) = \mathbf{0} \text{ and } \nabla^2\Phi(\mathbf{w}^*) \succ 0 \quad \Rightarrow \quad x^* \text{ strict local minimum}$$

$$\nabla\Phi(\mathbf{w}^*) = \mathbf{0} \text{ and } \nabla^2\Phi(\mathbf{w}^*) \prec 0 \quad \Rightarrow \quad x^* \text{ strict local maximum}$$

No conclusion can be drawn in the case $\nabla^2\Phi(\mathbf{w}^*)$ is indefinite!

Algebraic Characterization of **Unconstrained** Local Optima

Consider the unconstrained problem: $\min_{\mathbf{w}} \quad \Phi(\mathbf{w})$

1st-Order **Necessary** Condition of Optimality (FONC)

$$\mathbf{w}^* \text{ local optimum} \quad \Rightarrow \quad \nabla\Phi(\mathbf{w}^*) = \mathbf{0}, \;\; \mathbf{w}^* \text{ stationary point}$$

2nd-Order **Sufficient** Conditions of Optimality (SOSC)

NLP:

$$\nabla\Phi(\mathbf{w}^*) = \mathbf{0} \text{ and } \nabla^2\Phi(\mathbf{w}^*) \succ 0 \quad \Rightarrow \quad x^* \text{ strict local minimum}$$

$$\nabla\Phi(\mathbf{w}^*) = \mathbf{0} \text{ and } \nabla^2\Phi(\mathbf{w}^*) \prec 0 \quad \Rightarrow \quad x^* \text{ strict local maximum}$$

No conclusion can be drawn in the case $\nabla^2\Phi(\mathbf{w}^*)$ is indefinite!

Algebraic Characterization of **Unconstrained** Local Optima

Consider the unconstrained problem: $\min_{\mathbf{w}} \quad \Phi(\mathbf{w})$

1st-Order **Necessary** Condition of Optimality (FONC)

$$\mathbf{w}^* \text{ local optimum} \quad \Rightarrow \quad \nabla\Phi(\mathbf{w}^*) = \mathbf{0}, \;\; \mathbf{w}^* \text{ stationary point}$$

2nd-Order **Sufficient** Conditions of Optimality (SOSC)

NLP:

$$\nabla\Phi(\mathbf{w}^*) = \mathbf{0} \text{ and } \nabla^2\Phi(\mathbf{w}^*) \succ 0 \quad \Rightarrow \quad x^* \text{ strict local minimum}$$

$$\nabla\Phi(\mathbf{w}^*) = \mathbf{0} \text{ and } \nabla^2\Phi(\mathbf{w}^*) \prec 0 \quad \Rightarrow \quad x^* \text{ strict local maximum}$$

No conclusion can be drawn in the case $\nabla^2\Phi(\mathbf{w}^*)$ is indefinite!

Note:

- $\nabla\Phi(\mathbf{w}^*) = 0$ then $\nexists\,\mathbf{d}$ such that $\nabla\Phi(\mathbf{w}^*)^\top\mathbf{d} < 0$
- $\nabla^2\Phi \succ 0$ then $\forall\,\mathbf{d} \neq 0,\; \mathbf{d}^\top\nabla^2\Phi(\mathbf{w}^*)\mathbf{d} > 0$

Algebraic Characterization of **Unconstrained** Local Optima

Consider the unconstrained problem: $\min_{\mathbf{w}} \quad \Phi(\mathbf{w})$

1st-Order **Necessary** Condition of Optimality (FONC)

$$\mathbf{w}^* \text{ local optimum} \quad \Rightarrow \quad \nabla\Phi(\mathbf{w}^*) = \mathbf{0}, \;\; \mathbf{w}^* \text{ stationary point}$$

2nd-Order **Sufficient** Conditions of Optimality (SOSC)

NLP:

$$\nabla\Phi(\mathbf{w}^*) = \mathbf{0} \text{ and } \nabla^2\Phi(\mathbf{w}^*) \succ 0 \quad \Rightarrow \quad x^* \text{ strict local minimum}$$

$$\nabla\Phi(\mathbf{w}^*) = \mathbf{0} \text{ and } \nabla^2\Phi(\mathbf{w}^*) \prec 0 \quad \Rightarrow \quad x^* \text{ strict local maximum}$$

No conclusion can be drawn in the case $\nabla^2\Phi(\mathbf{w}^*)$ is indefinite!

Note:

- $\nabla\Phi(\mathbf{w}^*) = 0$ then $\nexists\, \mathbf{d}$ such that $\nabla\Phi(\mathbf{w}^*)^\top \mathbf{d} < 0$
- $\nabla^2\Phi \succ 0$ then $\forall\, \mathbf{d} \neq 0$, $\mathbf{d}^\top \nabla^2\Phi(\mathbf{w}^*)\mathbf{d} > 0$

Local optimum: "No *direction* $\mathbf{d}$ can improve the cost (locally)"

FONC for equality constraints

Consider the NLP problem:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}) = 0 \end{aligned}$$

FONC for equality constraints

Consider the NLP problem:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}) = 0 \end{aligned}$$

Definition: a point $\mathbf{w}$ satisfies LICQ[a] iff $\nabla \mathbf{g}(\mathbf{w})$ is full column rank

[a] Linear Independence Constraint Qualification

FONC for equality constraints

Consider the NLP problem:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}) = 0 \end{aligned}$$

Definition: a point $\mathbf{w}$ satisfies LICQ[a] iff $\nabla \mathbf{g}(\mathbf{w})$ is full column rank

[a] Linear Independence Constraint Qualification

First-order Necessary Conditions

Let $\Phi, \mathbf{g}$ in $\mathcal{C}^1$. If $\mathbf{w}^*$ is a (local) optimum, **and** $\mathbf{w}^*$ satisfies LICQ, then there is a unique vector $\boldsymbol{\lambda}$ such that:

$$\boxed{\begin{array}{lcr} \text{Dual feasibility:} & \nabla\Phi(\mathbf{w}^*) + \nabla\mathbf{g}(\mathbf{w}^*)\boldsymbol{\lambda} & = 0 \\ \text{Primal feasibility:} & \mathbf{g}(\mathbf{w}^*) & = 0 \end{array}}$$

FONC for equality constraints

Consider the NLP problem:

$$\min_{\mathbf{w}} \quad \Phi(\mathbf{w})$$
$$\text{s.t.} \quad \mathbf{g}(\mathbf{w}) = 0$$

Definition: a point $\mathbf{w}$ satisfies LICQ[a] iff $\nabla \mathbf{g}(\mathbf{w})$ is full column rank

[a] Linear Independence Constraint Qualification

Square system: $(n+m)$ conditions in $(n+m)$ variables $(\mathbf{w}, \boldsymbol{\lambda})$

Lagrange multipliers: $\boldsymbol{\lambda}_i \leftrightarrow \mathbf{g}_i$

Dual feasibility $\equiv$ **Lagrangian stationarity:**

$$\boxed{\nabla \mathcal{L}(\mathbf{w}^*, \boldsymbol{\lambda}^*) = \mathbf{0}}$$

where $\mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}) \triangleq \Phi(\mathbf{w}) + \boldsymbol{\lambda}^\top \mathbf{g}(\mathbf{w})$ is the Lagrangian

First-order Necessary Conditions

Let $\Phi, \mathbf{g}$ in $\mathcal{C}^1$. If $\mathbf{w}^*$ is a (local) optimum, **and** $\mathbf{w}^*$ satisfies LICQ, then there is a unique vector $\boldsymbol{\lambda}$ such that:

$$\boxed{\begin{array}{lcl} \text{Dual feasibility:} & \nabla \Phi(\mathbf{w}^*) + \nabla \mathbf{g}(\mathbf{w}^*)\boldsymbol{\lambda} & = 0 \\ \text{Primal feasibility:} & \mathbf{g}(\mathbf{w}^*) & = 0 \end{array}}$$

KKT point

Consider the NLP problem:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}) = 0 \\ & \mathbf{h}(\mathbf{w}) \leq 0 \end{aligned}$$

KKT point

Consider the NLP problem:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi\left(\mathbf{w}\right) \\ \text{s.t.} \quad & \mathbf{g}\left(\mathbf{w}\right) = 0 \\ & \mathbf{h}\left(\mathbf{w}\right) \leq 0 \end{aligned}$$

A point $(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ is called a KKT point if it satisfies:

where $\mathcal{L} = \Phi\left(\mathbf{w}\right) + \boldsymbol{\lambda}^T \mathbf{g}\left(\mathbf{w}\right) + \boldsymbol{\mu}^T \mathbf{h}\left(\mathbf{w}\right)$

KKT point

Consider the NLP problem:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}) = 0 \\ & \mathbf{h}(\mathbf{w}) \leq 0 \end{aligned}$$

A point $(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ is called a KKT point if it satisfies:

Dual Feasibility: $\nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*) = 0, \quad \boldsymbol{\mu}^* \geq 0,$

where $\mathcal{L} = \Phi(\mathbf{w}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{w}) + \boldsymbol{\mu}^T \mathbf{h}(\mathbf{w})$

KKT point

Consider the NLP problem:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}) = 0 \\ & \mathbf{h}(\mathbf{w}) \leq 0 \end{aligned}$$

A point $(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ is called a KKT point if it satisfies:

$$\boxed{\begin{array}{rl} \textbf{Dual Feasibility:} & \nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*) = 0, \quad \boldsymbol{\mu}^* \geq 0, \\ \textbf{Primal Feasibility:} & \mathbf{g}(\mathbf{w}^*) = 0, \quad \mathbf{h}(\mathbf{w}^*) \leq 0, \end{array}}$$

where $\mathcal{L} = \Phi(\mathbf{w}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{w}) + \boldsymbol{\mu}^T \mathbf{h}(\mathbf{w})$

KKT point

Consider the NLP problem:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}) = 0 \\ & \mathbf{h}(\mathbf{w}) \leq 0 \end{aligned}$$

A point $(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ is called a KKT point if it satisfies:

$$\boxed{\begin{array}{rl} \textbf{Dual Feasibility:} & \nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*) = 0, \quad \boldsymbol{\mu}^* \geq 0, \\ \textbf{Primal Feasibility:} & \mathbf{g}(\mathbf{w}^*) = 0, \quad \mathbf{h}(\mathbf{w}^*) \leq 0, \\ \textbf{Complementary Slackness:} & \boldsymbol{\mu}_i^* \mathbf{h}_i(\mathbf{w}^*) = 0, \quad \forall\, i \end{array}}$$

where $\mathcal{L} = \Phi(\mathbf{w}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{w}) + \boldsymbol{\mu}^T \mathbf{h}(\mathbf{w})$

First-Order Necessary Conditions (FONC)

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}) = 0 \\ & \mathbf{h}(\mathbf{w}) \le 0 \end{aligned}$$

First-Order Necessary Conditions

Let Φ, $\mathbf{g}$, $\mathbf{h}$ be $\mathcal{C}^1$. If $\mathbf{w}^*$ is a (local) optimum **and** satisfies LICQ, then there is a unique vector $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$ such that $(\mathbf{w}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ is a KKT point.

First-Order Necessary Conditions (FONC)

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}) = 0 \\ & \mathbf{h}(\mathbf{w}) \leq 0 \end{aligned}$$

First-Order Necessary Conditions

Let Φ, $\mathbf{g}$, $\mathbf{h}$ be $\mathcal{C}^1$. If $\mathbf{w}^*$ is a (local) optimum **and** satisfies LICQ, then there is a unique vector $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$ such that $(\mathbf{w}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ is a KKT point.

Active constraints:

- $\mathbf{h}_i(\mathbf{w}) < 0$ then $\boldsymbol{\mu}_i^* = 0$, and $\mathbf{h}_i$ is inactive
- $\boldsymbol{\mu}_i^* > 0$ and $\mathbf{h}_i(\mathbf{w}) = 0$ then $\mathbf{h}_i(\mathbf{w})$ is strictly active
- $\boldsymbol{\mu}_i^* = 0$ and $\mathbf{h}_i(\mathbf{w}) = 0$ then then $\mathbf{h}_i(\mathbf{w})$ is weakly active
- We define the **active set** $\mathbb{A}^*$ as the set of indices i of the active constraints

First-Order Necessary Conditions (FONC)

$$\begin{aligned} \min_{\mathbf{w}} \quad & \Phi(\mathbf{w}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{w}) = 0 \\ & \mathbf{h}(\mathbf{w}) \leq 0 \end{aligned}$$

Definition: a point $\mathbf{w}$ satisfies LICQ iff

$$\left[\nabla \mathbf{g}(\mathbf{w}), \quad \nabla \mathbf{h}_{\mathbb{A}^*}(\mathbf{w})\right]$$

is full column rank

First-Order Necessary Conditions

Let Φ, $\mathbf{g}$, $\mathbf{h}$ be $\mathcal{C}^1$. If $\mathbf{w}^*$ is a (local) optimum **and** satisfies LICQ, then there is a unique vector $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$ such that $(\mathbf{w}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ is a KKT point.

Active constraints:

- $\mathbf{h}_i(\mathbf{w}) < 0$ then $\boldsymbol{\mu}_i^* = 0$, and $\mathbf{h}_i$ is inactive
- $\boldsymbol{\mu}_i^* > 0$ and $\mathbf{h}_i(\mathbf{w}) = 0$ then $\mathbf{h}_i(\mathbf{w})$ is strictly active
- $\boldsymbol{\mu}_i^* = 0$ and $\mathbf{h}_i(\mathbf{w}) = 0$ then then $\mathbf{h}_i(\mathbf{w})$ is weakly active
- We define the **active set** $\mathbb{A}^*$ as the set of indices i of the active constraints

Outline

Some intuitions on the KKT conditions

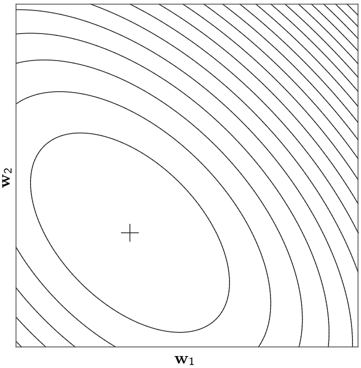
$$\min_{\mathbf{w}} \ \Phi(\mathbf{x})$$
$$\text{s.t. } h(\mathbf{w}) \leq \mathbf{0}$$

Ball rolling down a valley blocked by a fence

Some intuitions on the KKT conditions

$$\min_{\mathbf{w}} \; \Phi(\mathbf{x})$$
$$\text{s.t.} \; h(\mathbf{w}) \leq \mathbf{0}$$

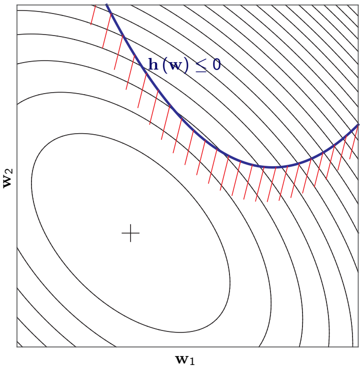
Ball rolling down a valley blocked by a fence



Some intuitions on the KKT conditions

$$\min_{\mathbf{w}} \ \Phi(\mathbf{x})$$
$$\text{s.t.} \ h(\mathbf{w}) \leq \mathbf{0}$$

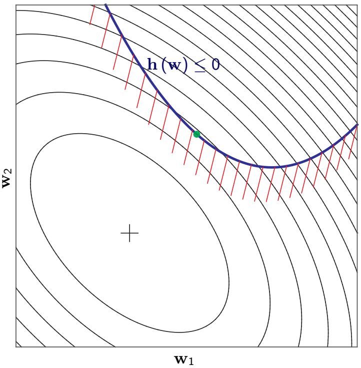
Ball rolling down a valley blocked by a fence



Some intuitions on the KKT conditions

$$\min_{\mathbf{w}} \; \Phi(\mathbf{x})$$
$$\text{s.t.} \; h(\mathbf{w}) \leq \mathbf{0}$$

Ball rolling down a valley blocked by a fence

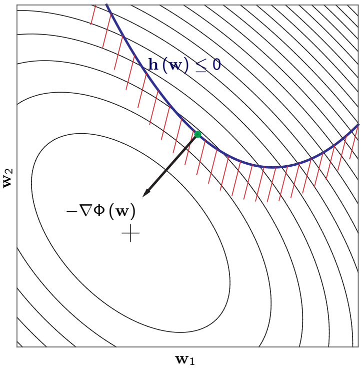


Some intuitions on the KKT conditions

$$\min_{\mathbf{w}} \ \Phi(\mathbf{x})$$
$$\text{s.t. } h(\mathbf{w}) \leq \mathbf{0}$$

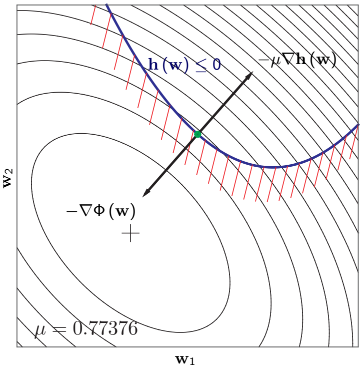
Ball rolling down a valley blocked by a fence

- $-\nabla\Phi$ is the gravity



Some intuitions on the KKT conditions

$$\min_{\mathbf{w}} \ \Phi(\mathbf{x})$$
$$\text{s.t.} \ h(\mathbf{w}) \leq \mathbf{0}$$



Ball rolling down a valley blocked by a fence

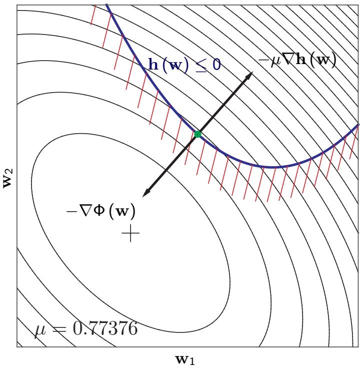
- $-\nabla\Phi$ is the gravity
- $-\mu\nabla h$ is the force of the fence. Sign $\mu \geq 0$ means the fence can only "push" the ball.

Some intuitions on the KKT conditions

$$\min_{\mathbf{w}} \ \Phi(\mathbf{x})$$
$$\text{s.t.} \ \ h(\mathbf{w}) \leq \mathbf{0}$$

Ball rolling down a valley blocked by a fence

- $-\nabla\Phi$ is the gravity
- $-\mu\nabla h$ is the force of the fence. Sign $\mu \geq 0$ means the fence can only "push" the ball.



Balance of the forces:

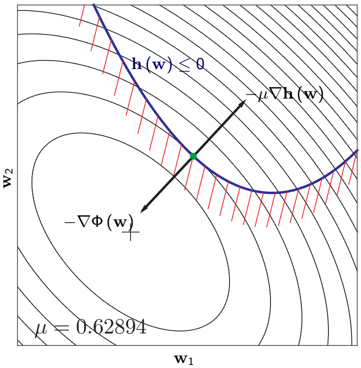
$$\nabla\mathcal{L} = \nabla\Phi(\mathbf{w}) + \mu\nabla h(\mathbf{w}) = 0$$

Some intuitions on the KKT conditions

$$\min_{\mathbf{w}} \; \Phi(\mathbf{x})$$
$$\text{s.t. } h(\mathbf{w}) \leq \mathbf{0}$$

Ball rolling down a valley blocked by a fence

- $-\nabla\Phi$ is the gravity
- $-\mu\nabla h$ is the force of the fence. Sign $\mu \geq 0$ means the fence can only "push" the ball.



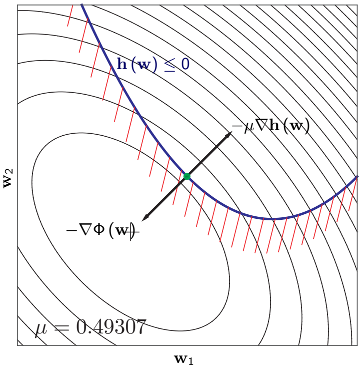
Balance of the forces:
$$\nabla\mathcal{L} = \nabla\Phi(\mathbf{w}) + \mu\nabla h(\mathbf{w}) = 0$$

Some intuitions on the KKT conditions

$$\min_{\mathbf{w}} \ \Phi(\mathbf{x})$$
$$\text{s.t. } h(\mathbf{w}) \leq \mathbf{0}$$

Ball rolling down a valley blocked by a fence

- $-\nabla\Phi$ is the gravity
- $-\mu\nabla h$ is the force of the fence. Sign $\mu \geq 0$ means the fence can only "push" the ball.



Balance of the forces:

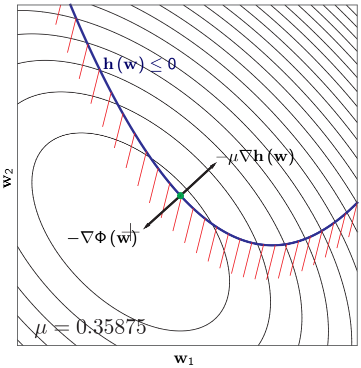
$$\nabla\mathcal{L} = \nabla\Phi(\mathbf{w}) + \mu\nabla h(\mathbf{w}) = 0$$

Some intuitions on the KKT conditions

$$\min_{\mathbf{w}} \ \Phi(\mathbf{x})$$
$$\text{s.t.} \ h(\mathbf{w}) \leq \mathbf{0}$$

Ball rolling down a valley blocked by a fence

- $-\nabla\Phi$ is the gravity
- $-\mu\nabla h$ is the force of the fence. Sign $\mu \geq 0$ means the fence can only "push" the ball.



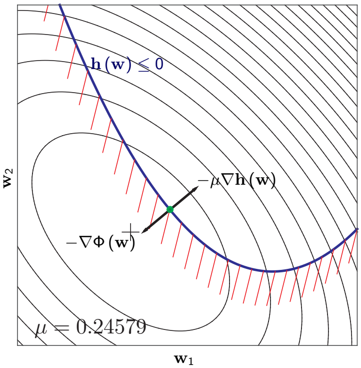
Balance of the forces:
$$\nabla\mathcal{L} = \nabla\Phi(\mathbf{w}) + \mu\nabla h(\mathbf{w}) = 0$$

Some intuitions on the KKT conditions

$$\min_{\mathbf{w}} \; \Phi(\mathbf{x})$$
$$\text{s.t.} \; h(\mathbf{w}) \leq \mathbf{0}$$

Ball rolling down a valley blocked by a fence

- $-\nabla\Phi$ is the gravity
- $-\mu\nabla h$ is the force of the fence. Sign $\mu \geq 0$ means the fence can only "push" the ball.



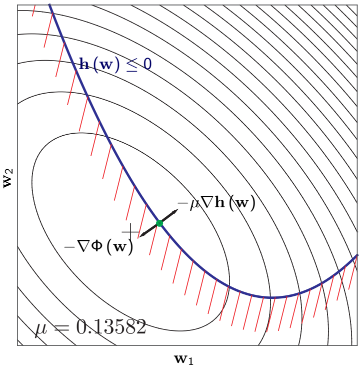
Balance of the forces:
$$\nabla\mathcal{L} = \nabla\Phi(\mathbf{w}) + \mu\nabla h(\mathbf{w}) = 0$$

Some intuitions on the KKT conditions

$$\min_{\mathbf{w}} \ \Phi(\mathbf{x})$$
$$\text{s.t.} \ h(\mathbf{w}) \leq \mathbf{0}$$

Ball rolling down a valley blocked by a fence

- $-\nabla\Phi$ is the gravity
- $-\mu\nabla h$ is the force of the fence. Sign $\mu \geq 0$ means the fence can only "push" the ball.



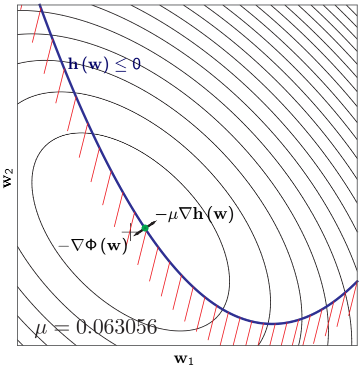
Balance of the forces:
$$\nabla\mathcal{L} = \nabla\Phi(\mathbf{w}) + \mu\nabla h(\mathbf{w}) = 0$$

Some intuitions on the KKT conditions

$$\min_{\mathbf{w}} \ \Phi(\mathbf{x})$$
$$\text{s.t.} \ h(\mathbf{w}) \leq \mathbf{0}$$

Ball rolling down a valley blocked by a fence

- $-\nabla\Phi$ is the gravity
- $-\mu\nabla h$ is the force of the fence. Sign $\mu \geq 0$ means the fence can only "push" the ball.



Balance of the forces:
$$\nabla\mathcal{L} = \nabla\Phi(\mathbf{w}) + \mu\nabla h(\mathbf{w}) = 0$$

Some intuitions on the KKT conditions

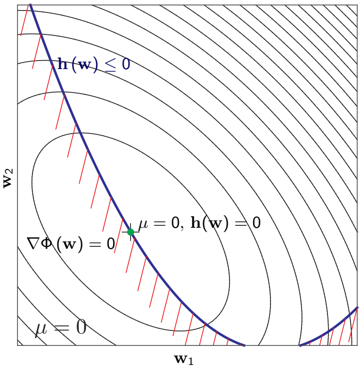
$$\min_{\mathbf{w}} \; \Phi(\mathbf{x})$$
$$\text{s.t. } h(\mathbf{w}) \leq \mathbf{0}$$

Ball rolling down a valley blocked by a fence

- $-\nabla\Phi$ is the gravity
- $-\mu\nabla h$ is the force of the fence. Sign $\mu \geq 0$ means the fence can only "push" the ball.
- Weakly active constraint:

$$h(\mathbf{w}) = 0, \quad \mu = 0$$

the ball touches the fence but no force is needed.



Balance of the forces:
$$\nabla\mathcal{L} = \nabla\Phi(\mathbf{w}) + \mu\nabla h(\mathbf{w}) = 0$$

Some intuitions on the KKT conditions

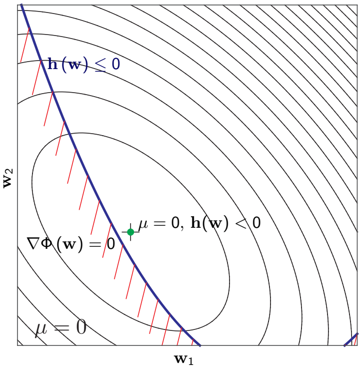
$$\min_{\mathbf{w}} \; \Phi(\mathbf{x})$$
$$\text{s.t.} \;\; h(\mathbf{w}) \leq \mathbf{0}$$

Ball rolling down a valley blocked by a fence

- $-\nabla\Phi$ is the gravity
- $-\mu\nabla h$ is the force of the fence. Sign $\mu \geq 0$ means the fence can only "push" the ball.
- Weakly active constraint:

$$h\left(\mathbf{w}\right) = 0, \quad \mu = 0$$

  the ball touches the fence but no force is needed.
- Inactive constraint $h\left(\mathbf{w}\right) < 0, \;\mu = 0$



Balance of the forces:
$$\nabla\mathcal{L} = \nabla\Phi\left(\mathbf{w}\right) + \mu\nabla h\left(\mathbf{w}\right) = 0$$

Some intuitions on the KKT conditions

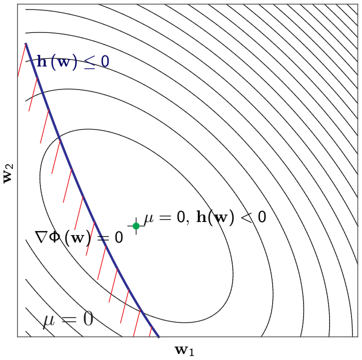
$$\min_{\mathbf{w}} \; \Phi(\mathbf{x})$$
$$\text{s.t. } h(\mathbf{w}) \leq \mathbf{0}$$

Ball rolling down a valley blocked by a fence

- $-\nabla\Phi$ is the gravity
- $-\mu\nabla h$ is the force of the fence. Sign $\mu \geq 0$ means the fence can only "push" the ball.
- Weakly active constraint:

  $$h(\mathbf{w}) = 0, \quad \mu = 0$$

  the ball touches the fence but no force is needed.
- Inactive constraint $h(\mathbf{w}) < 0, \; \mu = 0$
- Complementary slackness $\mu h = 0$ describes a contact problem (force exists only if the ball touches)



Balance of the forces:
$$\nabla\mathcal{L} = \nabla\Phi(\mathbf{w}) + \mu\nabla h(\mathbf{w}) = 0$$

Outline

1. KKT conditions

2. Some intuitions on the KKT conditions

3. Second Order Sufficient Conditions (SOSC)

Second-Order Sufficient Conditions for a Local Minimum

A point $(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ is called a KKT point if it satisfies:

Dual Feasibility:	$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*) = 0, \quad \boldsymbol{\mu}^* \geq 0,$
Primal Feasibility:	$\mathbf{g}(\mathbf{w}^*) = 0, \quad \mathbf{h}(\mathbf{w}^*) \leq 0,$
Complementary Slackness:	$\boldsymbol{\mu}_i^* \mathbf{h}_i(\mathbf{w}^*) = 0, \quad \forall\, i$

where $\mathcal{L} = \Phi(\mathbf{w}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{w}) + \boldsymbol{\mu}^T \mathbf{h}(\mathbf{w})$

Second-Order Sufficient Conditions for a Local Minimum

A point $(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ is called a KKT point if it satisfies:

Dual Feasibility:	$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*) = 0, \quad \boldsymbol{\mu}^* \geq 0,$
Primal Feasibility:	$\mathbf{g}(\mathbf{w}^*) = 0, \quad \mathbf{h}(\mathbf{w}^*) \leq 0,$
Complementary Slackness:	$\boldsymbol{\mu}_i^* \mathbf{h}_i(\mathbf{w}^*) = 0, \quad \forall\, i$

where $\mathcal{L} = \Phi(\mathbf{w}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{w}) + \boldsymbol{\mu}^T \mathbf{h}(\mathbf{w})$

- Let Φ, $\mathbf{g}$, $\mathbf{h}$ be $\mathcal{C}^2$

Second-Order Sufficient Conditions for a Local Minimum

A point $(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ is called a KKT point if it satisfies:

Dual Feasibility:	$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*) = 0, \quad \boldsymbol{\mu}^* \geq 0,$
Primal Feasibility:	$\mathbf{g}(\mathbf{w}^*) = 0, \quad \mathbf{h}(\mathbf{w}^*) \leq 0,$
Complementary Slackness:	$\boldsymbol{\mu}_i^* \mathbf{h}_i(\mathbf{w}^*) = 0, \quad \forall\, i$

where $\mathcal{L} = \Phi(\mathbf{w}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{w}) + \boldsymbol{\mu}^T \mathbf{h}(\mathbf{w})$

- Let Φ, $\mathbf{g}$, $\mathbf{h}$ be $\mathcal{C}^2$
- Suppose that $\mathbf{w}^*$ is regular and $\exists\ \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*$ such that $(\mathbf{w}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ is a KKT point

Second-Order Sufficient Conditions for a Local Minimum

A point $(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ is called a KKT point if it satisfies:

Dual Feasibility:	$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*) = 0, \quad \boldsymbol{\mu}^* \geq 0,$
Primal Feasibility:	$\mathbf{g}(\mathbf{w}^*) = 0, \quad \mathbf{h}(\mathbf{w}^*) \leq 0,$
Complementary Slackness:	$\boldsymbol{\mu}_i^* \mathbf{h}_i(\mathbf{w}^*) = 0, \quad \forall\, i$

where $\mathcal{L} = \Phi(\mathbf{w}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{w}) + \boldsymbol{\mu}^T \mathbf{h}(\mathbf{w})$

- Let Φ, $\mathbf{g}$, $\mathbf{h}$ be $\mathcal{C}^2$
- Suppose that $\mathbf{w}^*$ is regular and $\exists\, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*$ such that $(\mathbf{w}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ is a KKT point
- Set of feasible directions:

$$\mathcal{F} = \left\{ \mathbf{d} \quad | \quad \nabla \mathbf{g}(\mathbf{w}^*)^\top \mathbf{d} = 0, \quad \nabla \mathbf{h}_i(\mathbf{w}^*)^\top \mathbf{d} \leq 0, \quad \forall i \in \mathbb{A}^* \right\}$$

Second-Order Sufficient Conditions for a Local Minimum

A point $(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ is called a KKT point if it satisfies:

Dual Feasibility:	$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*) = 0, \quad \boldsymbol{\mu}^* \geq 0,$
Primal Feasibility:	$\mathbf{g}(\mathbf{w}^*) = 0, \quad \mathbf{h}(\mathbf{w}^*) \leq 0,$
Complementary Slackness:	$\boldsymbol{\mu}_i^* \mathbf{h}_i(\mathbf{w}^*) = 0, \quad \forall\, i$

where $\mathcal{L} = \Phi(\mathbf{w}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{w}) + \boldsymbol{\mu}^T \mathbf{h}(\mathbf{w})$

- Let Φ, $\mathbf{g}$, $\mathbf{h}$ be $\mathcal{C}^2$
- Suppose that $\mathbf{w}^*$ is regular and $\exists\ \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*$ such that $(\mathbf{w}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ is a KKT point
- Set of feasible directions:

$$\mathcal{F} = \left\{ \mathbf{d} \quad | \quad \nabla\mathbf{g}(\mathbf{w}^*)^\top \mathbf{d} = 0, \quad \nabla\mathbf{h}_i(\mathbf{w}^*)^\top \mathbf{d} \leq 0, \quad \forall i \in \mathbb{A}^* \right\}$$

- If for any $\mathbf{d} \in \mathcal{F} \setminus \{0\}$ with $\nabla\mathbf{h}_i(\mathbf{w}^*)^\top \mathbf{d} = 0$ for $\boldsymbol{\mu}_i^* > 0$ the inequality:

$$\mathbf{d}^\top \nabla^2 \mathcal{L}(\mathbf{w}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)\mathbf{d} > 0$$

  holds

Second-Order Sufficient Conditions for a Local Minimum

A point $(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ is called a KKT point if it satisfies:

Dual Feasibility:	$\nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*) = 0, \quad \boldsymbol{\mu}^* \geq 0,$
Primal Feasibility:	$\mathbf{g}(\mathbf{w}^*) = 0, \quad \mathbf{h}(\mathbf{w}^*) \leq 0,$
Complementary Slackness:	$\boldsymbol{\mu}_i^* \mathbf{h}_i(\mathbf{w}^*) = 0, \quad \forall\, i$

where $\mathcal{L} = \Phi(\mathbf{w}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{w}) + \boldsymbol{\mu}^T \mathbf{h}(\mathbf{w})$

- Let Φ, $\mathbf{g}$, $\mathbf{h}$ be $\mathcal{C}^2$
- Suppose that $\mathbf{w}^*$ is regular and $\exists\ \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*$ such that $(\mathbf{w}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ is a KKT point
- Set of feasible directions:

$$\mathcal{F} = \left\{ \mathbf{d} \quad \middle| \quad \nabla \mathbf{g}(\mathbf{w}^*)^\top \mathbf{d} = 0, \quad \nabla \mathbf{h}_i(\mathbf{w}^*)^\top \mathbf{d} \leq 0, \quad \forall i \in \mathbb{A}^* \right\}$$

- If for any $\mathbf{d} \in \mathcal{F} \setminus \{0\}$ with $\nabla \mathbf{h}_i(\mathbf{w}^*)^\top \mathbf{d} = 0$ for $\boldsymbol{\mu}_i^* > 0$ the inequality:

$$\mathbf{d}^\top \nabla^2 \mathcal{L}(\mathbf{w}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) \mathbf{d} > 0$$

  holds

$$\boxed{\textbf{Then, } \mathbf{w}^* \textbf{ is a local minimum.}}$$

Summary of Optimality Conditions

Optimality conditions for NLP with equality and/or inequality constraints:

- **1st-Order Necessary Conditions**: A regular local optimum of a (differentiable) NLP is a KKT point
- **2nd-Order Sufficient Conditions** require positivity of the Hessian in all critical feasible directions

Non-convex problem $\Rightarrow$ minimum is not necessarily global.
But some non-convex problems have a unique minimum.

Some important practical consequences...

- A local (global) optimum **may not** be a KKT point.
- A KKT point **may not** be a local (global) optimum.

... the lack of equivalence results from a lack of regularity and/or SOSC