Introduction to Optimization

Lecture 13: Constrained problems and optimality conditions.

From last class...

Linear programming and optimality conditions

We have

$$(LP) \qquad \min_{x\in\mathbb{R}^N} \{ c\cdot x \ : \ Ax + s = b, \ \text{and } s \geq 0 \}.$$

$$(DLP) \qquad \min_{y\in\mathbb{R}^M} \{ b\cdot y \ : \ A^T y + c = 0, \ \text{and } y \geq 0 \}.$$

Primal-dual optimality conditions

Primal feasibility: $Ax + s = b$ and $s \geq 0$.
Dual feasibility: $A^T y + c = 0$ and $y \geq 0$.
Complementarity: $y_i s_i = 0$ for $i = 1, \ldots, M$.

Idea of the proof

Write $K = N + M$ and $z = (x, s)^T$. Define $F : \mathbb{R}^K \to \mathbb{R} \cup \{+\infty\}$, $G : \mathbb{R}^M \to \mathbb{R} \cup \{+\infty\}$ and $P \in \mathbb{R}^{K \times M}$ by

$$F(z) = c \cdot x + \iota_{\mathbb{R}_+^M}(s), \quad G(y) = \iota_{\{b\}}(y) \quad \text{and} \quad Pz = Ax + s,$$

respectively, so that (LP) consists in minimizing $F(z) + G(Pz)$.

Idea of the proof

Write $K = N + M$ and $z = (x, s)^T$. Define $F : \mathbb{R}^K \to \mathbb{R} \cup \{+\infty\}$, $G : \mathbb{R}^M \to \mathbb{R} \cup \{+\infty\}$ and $P \in \mathbb{R}^{K \times M}$ by

$$F(z) = c \cdot x + \iota_{\mathbb{R}^M_+}(s), \quad G(y) = \iota_{\{b\}}(y) \quad \text{and} \quad Pz = Ax + s,$$

respectively, so that (LP) consists in minimizing $F(z) + G(Pz)$.We compute

$$\partial F(z) = \begin{bmatrix} c \\ \partial \iota_{\mathbb{R}^M_+}(s) \end{bmatrix}$$

Idea of the proof

Write $K = N + M$ and $z = (x, s)^T$. Define $F : \mathbb{R}^K \to \mathbb{R} \cup \{+\infty\}$, $G : \mathbb{R}^M \to \mathbb{R} \cup \{+\infty\}$ and $P \in \mathbb{R}^{K \times M}$ by

$$F(z) = c \cdot x + \iota_{\mathbb{R}^M_+}(s), \quad G(y) = \iota_{\{b\}}(y) \quad \text{and} \quad Pz = Ax + s,$$

respectively, so that (LP) consists in minimizing $F(z) + G(Pz)$.We compute

$$\partial F(z) = \begin{bmatrix} c \\ \partial \iota_{\mathbb{R}^M_+}(s) \end{bmatrix}, \quad \partial G(y) = \begin{cases} \emptyset & \text{if} \quad y \neq b \\ \mathbb{R}^M & \text{if} \quad y = b \end{cases}$$

Idea of the proof

Write $K = N + M$ and $z = (x, s)^T$. Define $F : \mathbb{R}^K \to \mathbb{R} \cup \{+\infty\}$, $G : \mathbb{R}^M \to \mathbb{R} \cup \{+\infty\}$ and $P \in \mathbb{R}^{K \times M}$ by

$$F(z) = c \cdot x + \iota_{\mathbb{R}^M_+}(s), \quad G(y) = \iota_{\{b\}}(y) \quad \text{and} \quad Pz = Ax + s,$$

respectively, so that (LP) consists in minimizing $F(z) + G(Pz)$.We compute

$$\partial F(z) = \begin{bmatrix} c \\ \partial \iota_{\mathbb{R}^M_+}(s) \end{bmatrix}, \quad \partial G(y) = \begin{cases} \emptyset & \text{if} \quad y \neq b \\ \mathbb{R}^M & \text{if} \quad y = b \end{cases}$$

and

$$P^T y = \begin{bmatrix} A^T y \\ y \end{bmatrix}.$$

Idea of the proof, continued

The primal-dual optimality conditions are

$$-P^T\hat{y} \in \partial F(\hat{z}) \quad \text{and} \quad \hat{y} \in \partial G(P\hat{x}),$$

which we translate as

$$\begin{aligned} -P^T\hat{y} \in \partial F(\hat{x}) \quad &\Leftrightarrow \quad -A^T\hat{y} = c, \quad \hat{s} \geq 0 \quad \text{and} \quad -\hat{y} \in \partial \iota_{\mathbb{R}^M_+}(\hat{s}). \\ \hat{y} \in \partial G(P\hat{x}) \quad &\Leftrightarrow \quad A\hat{x} + \hat{s} = b, \quad \hat{y} \in \mathbb{R}^M. \end{aligned}$$

Idea of the proof, continued

The primal-dual optimality conditions are

$$-P^T\hat{y} \in \partial F(\hat{z}) \quad \text{and} \quad \hat{y} \in \partial G(P\hat{x}),$$

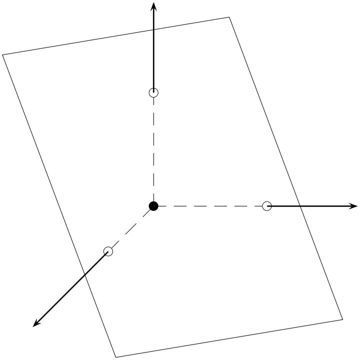
which we translate as

$$\begin{aligned} -P^T\hat{y} \in \partial F(\hat{x}) \quad &\Leftrightarrow \quad -A^T\hat{y} = c, \quad \hat{s} \geq 0 \quad \text{and} \quad -\hat{y} \in \partial \iota_{\mathbb{R}^M_+}(\hat{s}). \\ \hat{y} \in \partial G(P\hat{x}) \quad &\Leftrightarrow \quad A\hat{x} + \hat{s} = b, \quad \hat{y} \in \mathbb{R}^M. \end{aligned}$$

On the other hand, $\hat{y} \in \partial \iota_{\mathbb{R}^M_+}(\hat{s})$ means that, for $i = 1, \ldots, M$, $\hat{y}_i = 0$ if $\hat{s}_i > 0$, and $\hat{y}_i \geq 0$ if $\hat{s}_i = 0$.

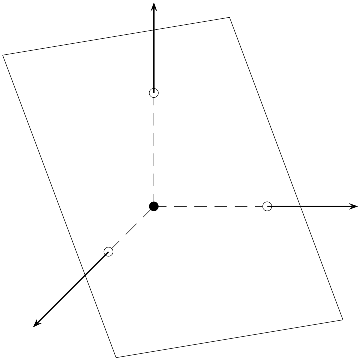
Example I

Distance from a plane to the origin



Example I

Distance from a plane to the origin

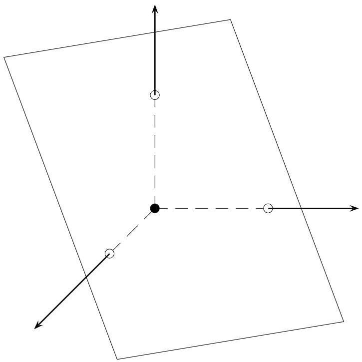


- Minimize distance to the origin

$$\sqrt{x^2 + y^2 + z^2}.$$

Example I

Distance from a plane to the origin



- Minimize distance to the origin

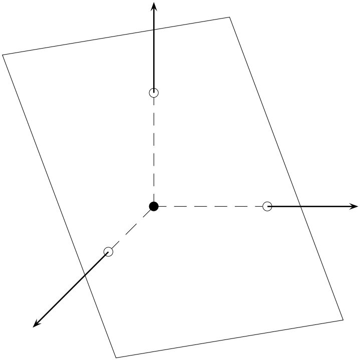
$$\sqrt{x^2 + y^2 + z^2}.$$

Equivalently, minimize

$$x^2 + y^2 + z^2.$$

Example I

Distance from a plane to the origin



- Minimize distance to the origin

$$\sqrt{x^2+y^2+z^2}.$$

  Equivalently, minimize

$$x^2+y^2+z^2.$$

- Subject to the constraint

$$ax+by+cz-d=0.$$

Optimization problems with equality constraints

Let $f, h_1, h_2, \ldots, h_M \in \mathcal{C}^1(\mathbb{R}^N; \mathbb{R})$, and consider the constrained problem

$$(\mathcal{P}) \qquad \min_{x \in C} f(x),$$

where the feasible set C is given by

$$C = \{x \in \mathbb{R}^N : h_m(x) = 0, \quad m = 1, \ldots, M\}.$$

Optimization problems with equality constraints

Let $f, h_1, h_2, \ldots, h_M \in \mathcal{C}^1(\mathbb{R}^N;\mathbb{R})$, and consider the constrained problem

$$(\mathcal{P}) \qquad \min_{x\in C} f(x),$$

where the feasible set C is given by

$$C = \{x \in \mathbb{R}^N : h_m(x) = 0, \quad m = 1, \ldots, M\}.$$

A local solution (constrained minimizer) $\hat{x}$ is regular if the set

$$V = \left\{\nabla h_1(\hat{x}), \nabla h_2(\hat{x}), \ldots, \nabla h_M(\hat{x})\right\}$$

is linearly independent.

Necessary conditions for optimality

Theorem (Lagrange Multiplier Theorem)

Let $\hat{x}$ be a regular local minimizer of $(\mathcal{P})$. There is a unique $\hat{\lambda} \in \mathbb{R}^M$, called Lagrange multiplier vector, such that

$$\nabla f(\hat{x}) + \sum_{m=1}^{M} \hat{\lambda}_m \nabla h_m(\hat{x}) = 0.$$

Necessary conditions for optimality

Theorem (Lagrange Multiplier Theorem)

Let $\hat{x}$ be a regular local minimizer of $(\mathcal{P})$. There is a unique $\hat{\lambda} \in \mathbb{R}^M$, called Lagrange multiplier vector, such that

$$\nabla f(\hat{x}) + \sum_{m=1}^{M} \hat{\lambda}_m \nabla h_m(\hat{x}) = 0.$$

If, in addition, $f, h_1, h_2, \ldots, h_M \in \mathcal{C}^2(\mathbb{R}^N;\mathbb{R})$, then

$$y \cdot \left(\nabla^2 f(\hat{x}) + \sum_{m=1}^{M} \hat{\lambda}_m \nabla^2 h_m(\hat{x}) \right) y \geq 0$$

for all $y \in V^\perp$.

Sketch of the proof

- Pick x_k that minimizes

$$f_k(x) = f(x) + \frac{\varepsilon}{2}\|x - \hat{x}\|^2 + \frac{k}{2}\sum_{m=1}^{M} \|h_m(x)\|^2$$

over a small ball around $\hat{x}$.

Sketch of the proof

- Pick x_k that minimizes

$$f_k(x) = f(x) + \frac{\varepsilon}{2}\|x - \hat{x}\|^2 + \frac{k}{2}\sum_{m=1}^{M} \|h_m(x)\|^2$$

over a small ball around $\hat{x}$.

- Let $k \to \infty$ and and use the optimality condition $\nabla f_k(x_k) = 0$ to deduce that $x_k \to \hat{x}$ and $kh_m(x_k) \to s_m$.

Sketch of the proof

- Pick x_k that minimizes

$$f_k(x) = f(x) + \frac{\varepsilon}{2}\|x - \hat{x}\|^2 + \frac{k}{2}\sum_{m=1}^{M} \|h_m(x)\|^2$$

  over a small ball around $\hat{x}$.

- Let $k \to \infty$ and and use the optimality condition $\nabla f_k(x_k) = 0$ to deduce that $x_k \to \hat{x}$ and $kh_m(x_k) \to s_m$.

- Show that $\hat{\lambda}_m = s_m$ has the desired property.

Sketch of the proof

- Pick x_k that minimizes

$$f_k(x) = f(x) + \frac{\varepsilon}{2}\|x - \hat{x}\|^2 + \frac{k}{2}\sum_{m=1}^{M} \|h_m(x)\|^2$$

  over a small ball around $\hat{x}$.

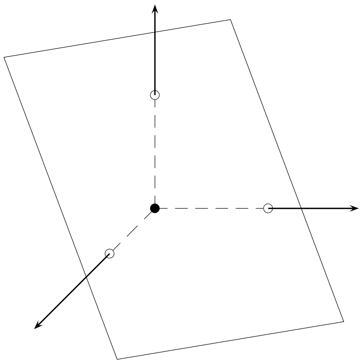
- Let $k \to \infty$ and and use the optimality condition $\nabla f_k(x_k) = 0$ to deduce that $x_k \to \hat{x}$ and $kh_m(x_k) \to s_m$.

- Show that $\hat{\lambda}_m = s_m$ has the desired property.

- Proceed analogously to obtain the second order condition.

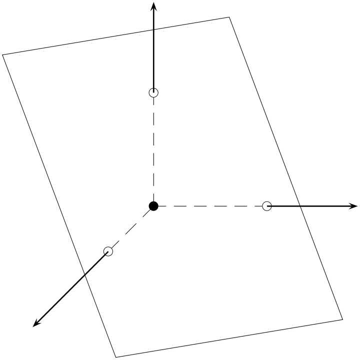
Back to Example I

Find the point in the plane $ax + by + cz = d$ that is closest to the origin



Back to Example I

Find the point in the plane $ax + by + cz = d$ that is closest to the origin

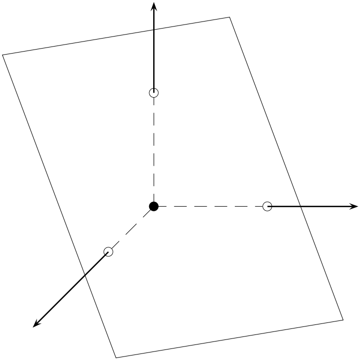


- Minimize distance to the origin

$$\sqrt{x^2 + y^2 + z^2}.$$

Back to Example I

Find the point in the plane $ax + by + cz = d$ that is closest to the origin



- Minimize distance to the origin

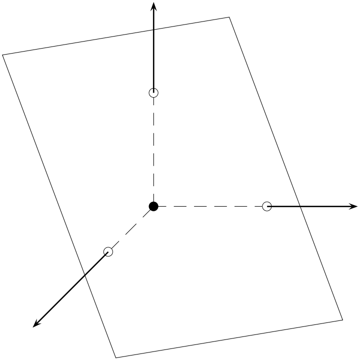
$$\sqrt{x^2 + y^2 + z^2}.$$

Equivalently, minimize

$$x^2 + y^2 + z^2.$$

Back to Example I

Find the point in the plane $ax + by + cz = d$ that is closest to the origin



- Minimize distance to the origin

$$\sqrt{x^2 + y^2 + z^2}.$$

Equivalently, minimize

$$x^2 + y^2 + z^2.$$

- Subject to the constraint

$$ax + by + cz - d = 0.$$

Second order sufficient conditions

Theorem

Let $f, h_1, h_2, \ldots, h_M : \mathbb{R}^N \to \mathbb{R}$ be twice continuously differentiable, and let $\hat{x} \in C$ and $\hat{\lambda} \in \mathbb{R}^M$ satisfy

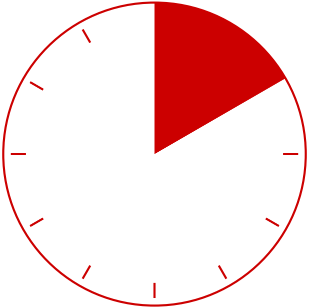
$$\nabla f(\hat{x}) + \sum_{m=1}^{M} \hat{\lambda}_m \nabla h_m(\hat{x}) = 0$$

and

$$y \cdot \left(\nabla^2 f(\hat{x}) + \sum_{m=1}^{M} \hat{\lambda}_m \nabla^2 h_m(\hat{x}) \right) y > 0$$

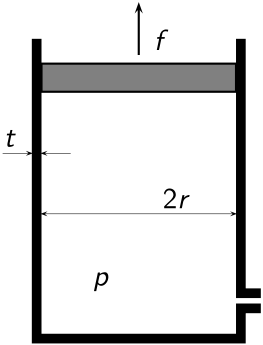
for all $y \in V^\perp \setminus \{0\}$. Then, $\hat{x}$ is a strict local minimizer of $(\mathcal{P})$.

Break

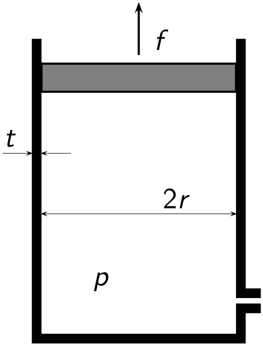


Example II

Hydraulic Cylinder Design



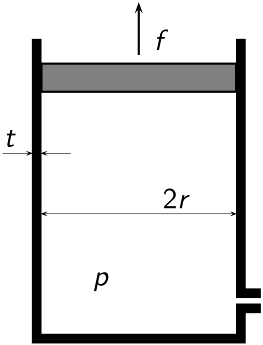
- Minimal force required
  - $f = \pi r^2 p \geq f_{min}$.



- Minimal force required
  - $f = \pi r^2 p \geq f_{min}$.
- Maximal hoop stress
  - $s = \dfrac{pr}{t} \leq s_{max}$

Example II

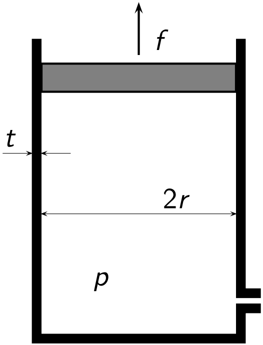
Hydraulic Cylinder Design – Manufacturing Constraints



- Minimal force required
  - $f = \pi r^2 p \geq f_{min}$.
- Maximal hoop stress
  - $s = \dfrac{pr}{t} \leq s_{max}$
- Minimal thickness
  - $t \geq t_{min}$

Example II

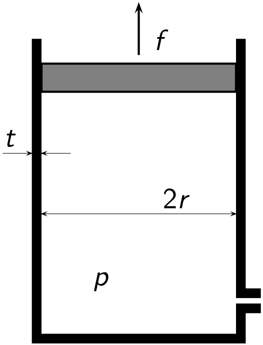
Hydraulic Cylinder Design – Manufacturing Constraints



- Minimal force required
    - $f = \pi r^2 p \geq f_{min}$.
- Maximal hoop stress
    - $s = \dfrac{pr}{t} \leq s_{max}$
- Minimal thickness
    - $t \geq t_{min}$
- Maximal pressure
    - $p \leq p_{max}$.

Example II

Hydraulic Cylinder Design – Optimization Problem



- Minimize the total width $2(r+t)$
- Subject to the constraints
  - $pr - ts_{max} \leq 0$
  - $f_{min} - \pi r^2 p \leq 0$
  - $t_{min} - t \leq 0$
  - $p - p_{max} \leq 0.$

Inequality constraints

Let $f, g_j, h_m \in \mathcal{C}^1(\mathbb{R}^N;\mathbb{R})$, and consider the constrained problem

$$(\mathcal{P}) \qquad \min_{x \in C} f(x),$$

where the feasible set C is now given by

$$C = \{\, x \in \mathbb{R}^N \ : \ h_m(x) = 0,\ g_j(x) \leq 0,\ \forall m, j \,\}.$$

Inequality constraints

Let $f, g_j, h_m \in \mathcal{C}^1(\mathbb{R}^N;\mathbb{R})$, and consider the constrained problem

$$(\mathcal{P}) \qquad \min_{x\in C} f(x),$$

where the feasible set C is now given by

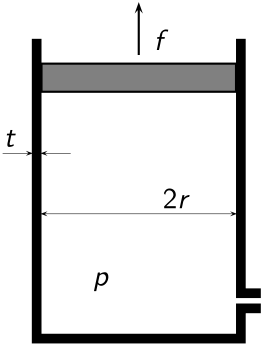
$$C = \{\, x \in \mathbb{R}^N \;:\; h_m(x) = 0,\; g_j(x) \leq 0,\; \forall m, j \,\}.$$

The set of active inequality constraints at a local solution $\hat{x}$ is

$$A(\hat{x}) = \{\, j \;:\; g_j(\hat{x}) = 0 \,\}.$$

Back to Example II

Hydraulic Cylinder Design – Optimization Problem



- Minimize the total width $2(r+t)$
- Subject to the constraints
  - $pr - ts_{max} \leq 0$
  - $f_{min} - \pi r^2 p \leq 0$
  - $t_{min} - t \leq 0$
  - $p - p_{max} \leq 0.$

Necessary conditions for optimality

Theorem (Karush-Kuhn-Tucker conditions)

Let $\hat{x}$ be a regular local minimizer of $(\mathcal{P})$. There exist unique Lagrange multiplier vectors $\hat{\lambda} \in \mathbb{R}^M$ and $\hat{\mu} \in \mathbb{R}^J_+$, such that

$$\nabla f(\hat{x}) + \sum_{j=1}^{J} \hat{\mu}_j \nabla g_j(\hat{x}) + \sum_{m=1}^{M} \hat{\lambda}_m \nabla h_m(\hat{x}) = 0,$$

and $\hat{\mu}_j = 0$ for all $j \notin A(\hat{x})$, which means that $\hat{\mu}_j g_j(\hat{x}) = 0$ for all j.

Necessary conditions for optimality

Theorem (Karush-Kuhn-Tucker conditions)

Let $\hat{x}$ *be a regular local minimizer of* $(\mathcal{P})$. *There exist unique Lagrange multiplier vectors* $\hat{\lambda} \in \mathbb{R}^M$ *and* $\hat{\mu} \in \mathbb{R}^J_+$, *such that*

$$\nabla f(\hat{x}) + \sum_{j=1}^{J} \hat{\mu}_j \nabla g_j(\hat{x}) + \sum_{m=1}^{M} \hat{\lambda}_m \nabla h_m(\hat{x}) = 0,$$

and $\hat{\mu}_j = 0$ *for all* $j \notin A(\hat{x})$, *which means that* $\hat{\mu}_j g_j(\hat{x}) = 0$ *for all* j.

If, in addition, $f, g_j, h_m \in \mathcal{C}^2(\mathbb{R}^N;\mathbb{R})$, *then*

$$y \cdot \left(\nabla^2 f(\hat{x}) + \sum_{m=1}^{M} \hat{\lambda}_m \nabla^2 h_m(\hat{x}) \right) y \geq 0$$

for all $y \in \mathcal{V}^\perp$, *where* $\mathcal{V} = \{\nabla g_j(\hat{x}), \nabla h_m(\hat{x}),\ j \in A(\hat{x}),\ m = 1, \ldots M\}$.