2.3 Problems under general constraints

(1) Conditions for a minimum under constraints and the projection thm

Thm. (Euler's inequalities)

V normed vector space $U \subset V$ convex, nonempty

$F : \mathcal{O}(U) \to F$ differentiable at $\hat{u} \in U$.

(i) if F has at $\hat{u}$ a minimum constrained to U

then $\quad \langle \nabla F(\hat{u}), u - \hat{u} \rangle \geq 0 \quad \forall u \in U \quad (*)$

(ii) if in addition F is convex :

F has at $\hat{u}$ a min. const. to U if and only if

$(*)$ holds true

(iii) If U is open $\quad (*) \iff \nabla F(\hat{u}) = 0$

<u>The projection thm.</u>

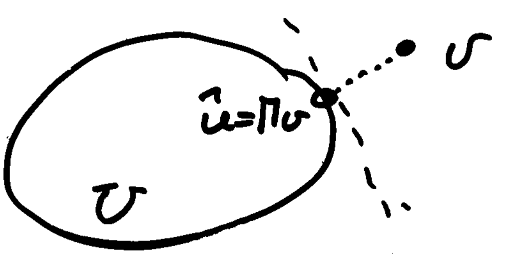
V Hilbert, real $\quad U \subset V$ nonempty convex closed

If $v \in V$, then (i) $\iff$ (ii)

(i) $\exists\, \hat{u} = \underline{\Pi v} \in U$ unique and satisfies

$$\|v - \hat{u}\| = \inf_{u \in U} \|v - u\|$$

(ii) $\hat{u} \in U, \ \langle v - \hat{u}, u - \hat{u} \rangle \leq 0 \quad \forall u \in U$



$\hat{u}$ is the projection of v on U

Proof. $F(u) = \frac{1}{2}\|v - u\|^2$

continuous, convex, coercive ...

$\inf_{u \in U} F(u)$ – this pb has a solution

convex

$$\boxed{\exists\, \hat{u} : \quad F(\hat{u}) = \inf_{u \in U} F(u)}$$

By Euler:

$\nabla F_{(u)} = u - v \qquad \left(F(u) = \frac{1}{2}\langle u - v, u - v \rangle \right)$

$\nabla^2 F = I$

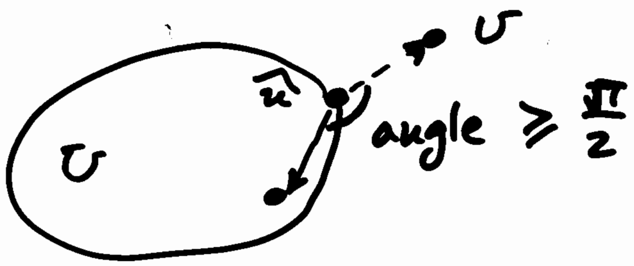
$\Rightarrow$ F strictly convex, elliptic $\Rightarrow$ unique min. + strict.

so we proved (i). $\qquad \|v - \hat{u}\| \leq \|v - u\| \quad \forall u \in U$

Euler to exhibit (ii). $\hat{u}$ satisfies (i) $\Leftrightarrow$

$$\langle \underbrace{\nabla F(\hat{u})}_{\hat{u} - v}, u - \hat{u} \rangle \geq 0 \quad \forall u \in U$$

$$\langle v - \hat{u}, u - \hat{u} \rangle \leq 0 \quad \forall u \in U$$



Thm.

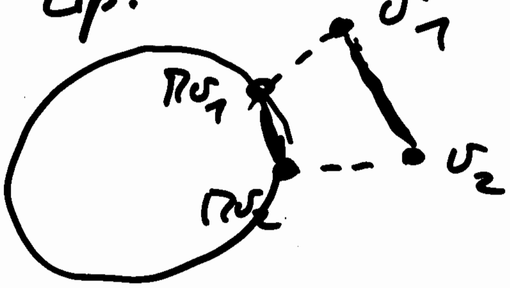
Π is called the projection operator.

Properties of $\Pi : V \to U$

(i) $v - \Pi v = 0 \iff v \in U$

→ (ii) $\|\Pi v_1 - \Pi v_2\| \le \|v_1 - v_2\| \quad \forall v_1, v_2 \in V$

$\Rightarrow \Pi$ unif. continuous Lip.



(iii) Π is linear $\iff$
U is a vector subspace

In such a case

$$v - \Pi v \perp u, \quad \forall u \in U$$

Proof of (ii)

$$\underline{\|\Pi v_1 - \Pi v_2\|^2} = \langle \Pi v_1 - \Pi v_2, \Pi v_1 - \Pi v_2 - v_1 + (v_1 - v_2) + v_2 \rangle$$

$$= \langle \Pi v_1 - \Pi v_2, \Pi v_1 - v_1 \rangle \le 0 \text{ by proj thm.}$$

$$+ \langle \Pi v_1 - \Pi v_2, v_1 - v_2 \rangle$$

$$+ \langle \Pi v_1 - \Pi v_2, v_2 - \Pi v_2 \rangle \le 0 \text{ by proj thm.}$$

$$\le \underline{\|\Pi v_1 - \Pi v_2\|} \cdot \|v_1 - v_2\| \quad \text{(Schwarz)} \Rightarrow \text{(ii)}$$

if U vector subspace

$$\langle v - \hat{u}, u - \hat{u} \rangle \le 0 \qquad \forall u$$

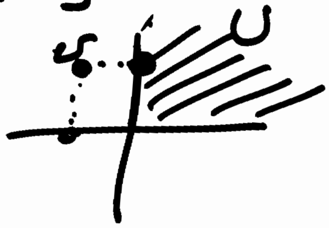
$$\uparrow \; 2\hat{u} - u \in U$$

$$\langle v - \hat{u}, \hat{u} - u \rangle \le 0 \qquad \Rightarrow \; = 0.$$

Examples

$$U = \{ u \in \mathbb{R}^n : u[i] \geq 0 \quad 1 \leq i \leq n \}$$

$$\Pi v[i] = \max(v[i], 0)$$



$$U = \{ u \in \mathbb{R}^n \quad a_i \leq u[i] \leq b_i \} \quad (*)$$

$$\Pi v(i) = \begin{cases} a_i & \text{if } v(i) < a_i \\ v(i) & \text{if } a_i \leq v(i) \leq b_i \\ b_i & \text{if } v(i) > b_i \end{cases}$$

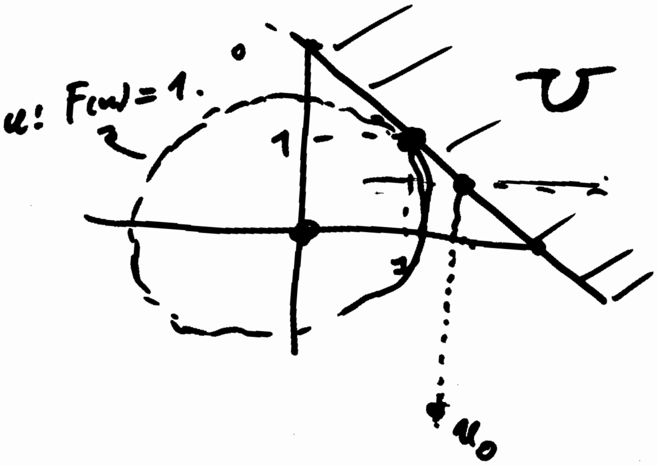
(2) The relaxation method can be applied when U is as above.

$$\left. \begin{aligned} & F(u_{k+1}(1), \dots u_{k+1}(i), \dots u_k(n)) = \\ & = \inf_{a_i \leq \rho \leq b_i} F(u_{k+1}(1), \dots, \rho, \dots u_k(n)) \end{aligned} \right] u_k$$

Cannot be extended to more general constraints

$$F(u) = u(1)^2 + u(2)^2$$

$$U = \{ u \in \mathbb{R}^2 : u(1) + u(2) \geq 2 \} \qquad u_0 \neq 1$$



Thm F elliptic and U is of the form $(*)$ then u_k converges.

2.3.3. Projected gradient with a variable step.

U convex closed $\quad F$ convex

Motivation.

$\hat{u} \in U$ and $F(\hat{u}) = \inf\limits_{u \in U} F(u)$

$\iff$ $\hat{u} \in U$ and $\langle \rho \nabla F(\hat{u}), u - \hat{u} \rangle \geq 0 \quad \forall u \in U,\ \rho > 0$

$\iff$ $\hat{u} \in U$ and $\langle -\rho \nabla F(\hat{u}), u - \hat{u} \rangle \leq 0 \quad \forall u \in U,\ \rho > 0$

$\hat{u} - \hat{u}$

$\iff$ $\hat{u} \in U$ and $\langle \underbrace{\hat{u} - \rho \nabla F(\hat{u})}_{v} - \hat{u}, u - \hat{u} \rangle \leq 0 \quad \forall u \in U,\ \rho > 0$

$\hat{u}$ is the projection of v!

$$\hat{u} = \Pi(\hat{u} - \rho \nabla F(\hat{u})) \quad \rho > 0$$

$\Rightarrow$ $\hat{u}$ is the fixed point of $G : V \to V$ where

$$G(u) = \Pi(u - \rho \nabla F(u))$$

we will use the method of successive projections:

$$u_{k+1} = G(u_k) \qquad k \in \mathbb{N}$$

If $U = V$ then Π = identity ... the grad. method.

G is a <u>contraction</u> if $\exists\ \gamma \in (0,1)$

$\|G(u) - G(v)\| \leq \gamma \|u - v\| \quad \forall u, v \in V$

The fixed pt then says that $\exists!\ \hat{u} : G(\hat{u}) = \hat{u}$

(V complete)

We will choose ρ such that G is a contraction.

$$u_{k+1} = \Pi \left(u_k - \rho_k \nabla F(u_k) \right) \quad \rho_k > 0 \quad k \geq 0$$

Thm $F : V \to \mathbb{R}$ differentiable. Suppose that $\exists\, \mu > 0$ and $\exists\, M > 0$ such that

① $\langle \nabla F(u_1) - \nabla F(u_2), u_1 - u_2 \rangle \geq \mu \| u_1 - u_2 \|^2 \quad \forall u_1, u_2 \in V$

② $\| \nabla F(u_1) - \nabla F(u_2) \| \leq M \| u_1 - u_2 \| \quad \forall u_1, u_2 \in V$

Let $\zeta \in (0, \frac{\mu}{M^2})$

If $\zeta < \rho_k \leq \frac{2\mu}{M} - \zeta \quad \forall k \in \mathbb{N}$

then (u_k) converges et

$$\| u_k - \hat{u} \| \leq \gamma^k \| u_0 - \hat{u} \|$$

$$\gamma = \sqrt{1 - 2\mu\zeta + M^2 \zeta^2} < 1.$$

Proof. $G_k(u) = \Pi \left(u - \rho_k \nabla F(u) \right) \in U$

$$u_{k+1} = G_k(u_k)$$

we have to show that G_k is a contraction $\forall k$.

$u_1, u_2 \in V$ arbitrary:

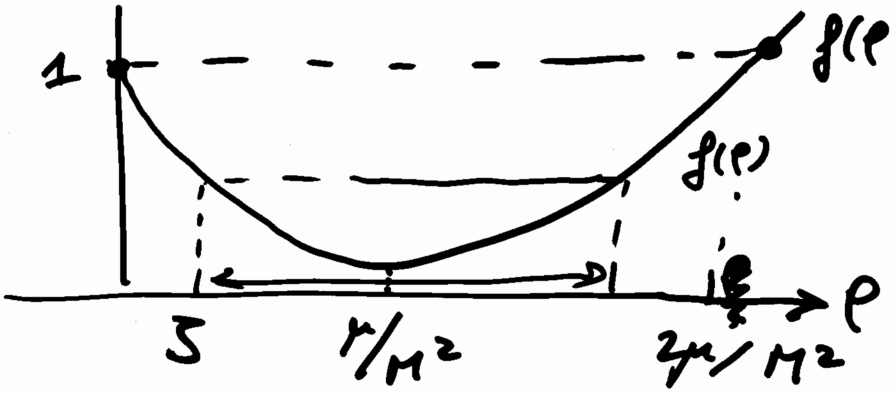
$$\underline{\| G_k(u_1) - G_k(u_2) \|^2} = \| \Pi (u_1 - \rho_k \nabla F(u_1)) - \Pi (u_2 - \rho_k \nabla F(u_2)) \|^2$$

$$\leq \| u_1 - u_2 - \rho_k (\nabla F(u_1) - \nabla F(u_2)) \|^2$$

$$= \| u_1 - u_2 \|^2 - 2\rho_k \overbrace{\langle \nabla F(u_1) - \nabla F(u_2), u_1 - u_2 \rangle}^{\text{apply ①}} + \rho_k^2 \overbrace{\| \nabla F(u_1) - \nabla F(u_2) \|^2}^{\text{apply ②}}$$

$$\leq \| u_1 - u_2 \|^2 - 2\rho_k \mu \| u_1 - u_2 \|^2 + \rho_k^2 M^2 \| u_1 - u_2 \|^2$$

$$= \underbrace{(1 - 2\mu\rho_k + \rho_k^2 M^2)}_{\text{must be } < 1 \text{ (and } > 0)} \| u_1 - u_2 \|^2$$



$\|u_{k+1}-\hat{u}_k\| \le \gamma \|u_k-\hat{u}\| \dots \le \gamma^{k+1}\|u_0-\hat{u}\|.$

The gradiant method with a variable step is a particular case - hence we have been the proof.

Rmk. the cond. on ρ_k is sufficient.

Gradient with projection for an elliptic F.

$F(u)=\frac{1}{2}\langle Bu,u\rangle - \langle c,u\rangle$

~~Ex~~ $u_{k+1} = u_k - \rho_k (Bu_k - c)$

$\nabla F(u) = Bu - C$

$\|\nabla F(u_1) - \nabla F(u_2)\| = \|B(u_1-u_2)\| \ge \overbrace{\lambda_{min}(B)}^{\mu} \cdot \|u_1-u_2\|$

$\|\nabla F(u_1) - \nabla F(u_2)\| \le \underbrace{\lambda_{max}(B)}_{M} \|u_1-u_2\|$

The thm recommends $\zeta \le \rho_k \le \frac{2\lambda_{min}}{\lambda^2_{max}} - \zeta$

$\zeta \in \left(0, \frac{\lambda_{min}}{\lambda^2_{max}}\right)$

We can improve

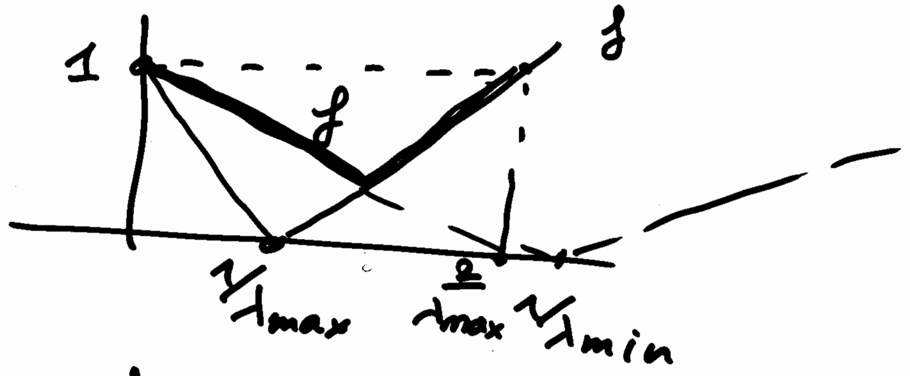
$u_{k+1}-\hat{u} = u_k - \hat{u} - \rho_k B(u_k-\hat{u})$

$= (I-\rho_k B)(u_k-\hat{u})$

$\|u_{k+1}-\hat{u}\| \le \|\underbrace{I-\rho_k B}_{\text{symmetric}}\|_2 \|u_k-\hat{u}\|$

its spectral radius = the max. eigenvalue

$$f(\rho) = \max \left\{ |1-\rho_* \lambda_{min}(B)|, |1-\rho_* \lambda_{max}(B)| \right\}$$



$\rightarrow$ it is hence enough $\rho < \frac{2}{\lambda_{max}}$

2.3.4 Penalty method.

Replace the constrained optim. pb by a family of unconstrained pbs.

We define (construct) $G : \mathbb{R}^n \to \mathbb{R}$
continuous, convex $G(u) \geq 0 \quad \forall u \in \mathbb{R}^n$
and $G(u) = 0 \Leftrightarrow u \in U$

For every $\varepsilon > 0$

$$\mathcal{F}_\varepsilon(u) = F(u) + \frac{1}{\varepsilon} G(u)$$

we find the solution $\hat{u}$ by letting $\varepsilon \to 0$

Thm. $F : \mathbb{R}^n \to \mathbb{R}$ continuous coercive strictly convex
U convex, defined by G. Then

(i) $\forall \varepsilon > 0 \quad \exists u_\varepsilon$ unique sth $\mathcal{F}_\varepsilon(u_\varepsilon) = \inf_{u \in \mathbb{R}^n} \mathcal{F}_\varepsilon(u)$

(ii) $\lim_{\varepsilon \to 0} u_\varepsilon = \hat{u}$ (the unique solution of the original pb.)

Proof. $\mathcal{F}_\varepsilon(u) \geq F(u) \Rightarrow \mathcal{F}_\varepsilon$ coercive;
strictly convex $\Rightarrow$ (i)

$$\mathbf{F}(\overset{\in U}{u_\varepsilon}) \le \underbrace{F(u_\varepsilon) + \tfrac{1}{\varepsilon} G(u_\varepsilon)} = \mathcal{F}_\varepsilon(u_\varepsilon)$$

$$\le \mathcal{F}_\varepsilon(\hat{u}) = \underline{F(\hat{u})}$$

all. $\mathcal{F}_\varepsilon$ coercive ... $(u_\varepsilon)_{\varepsilon > 0}$ is bounded.

$\Rightarrow$ convergent subsequence $(u_{\varepsilon_k})_{k \ge 0}$

$$u_{\varepsilon_k} \xrightarrow[k\to\infty]{} \tilde{u}$$

$F(u_{\varepsilon_k}) \le F(\hat{u})$ / $F(\tilde{u}) = \lim_{k\to\infty} \mathbf{F}(u_{\varepsilon_k}) \le F(\hat{u})$

F continuous

$$0 \le G(u_{\varepsilon_k}) \le \varepsilon_k \big(\underbrace{F(\hat{u}) - F(u_{\varepsilon_k})}_{\le \text{ constant}}\big)$$

$\uparrow$ goes to 0

$$\rightarrow \lim_{k\to\infty} G(u_{\varepsilon_k}) = 0 = G(\tilde{u})$$

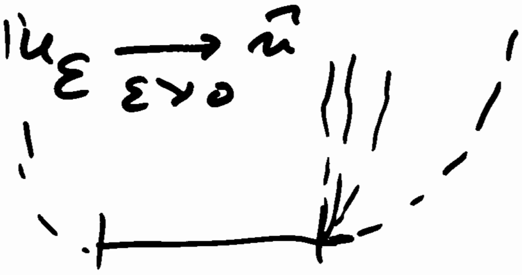
$\Rightarrow \tilde{u} \in U$

$F(\tilde{u}) \le F(\hat{u})$

we have a unique solution

$\tilde{u} = \hat{u}$

$u_\varepsilon \xrightarrow[\varepsilon \to 0]{} \hat{u}$



No systematic way to construct G.

G may not be diff. or conv. — additional difficulty

2.4 Pbs with equality constraints

2.4.1 Lagrange multipliers.

Thm. (Necessary cond. for a constrained min)

$V = V_1 \times V_2$ normed vector space

V_2 is complete, $\mathcal{O} \subset V$ open.

$G : \mathcal{O} \to V_2$, C^1

$\to U = \{ u \in \mathcal{O} : G(u_1, \underline{u_2}) = 0 \}$

$\hat{u} = (\hat{u}_1, \hat{u}_2)$ solution

$\underline{D_2 G(\hat{u}_1, \hat{u}_2) \in \text{Isomorphism}(V_2)}$

F differentiable at $\hat{u}$

$\Rightarrow \exists \lambda(\hat{u}) \in \mathcal{L}(V_2, \mathbb{R})$ such that

$$DF(\hat{u}) + \lambda(\hat{u}) \, DG(\hat{u}) = 0$$

Implicit func. thm. $\hat{u} = (\hat{u}_1, \hat{u}_2)$

$\exists \mathcal{O}_1 \subset V_1$, $\mathcal{O}_2 \subset V_2$ $f : \mathcal{O}_1 \to \mathcal{O}_2$ such that

$\mathcal{O}_1 \times \mathcal{O}_2 \cap U = \{ (u_1, u_2) : u_2 = f(u_1) \}$

and $G(u_1, f(u_1)) = 0 \quad \forall u_1 \in \mathcal{O}_1$

$\cancel{0 = D_1 G f(u_1)} + D_2$

$$0 = D_1 G(u_1, f(u_1)) + D_2 G(u_1, f(u_1)) \, Df(u_1)$$

$$Df(u_1) = - \underline{D_2 G(u_1, f(u_1))}^{-1} D_1 G(u_1, f(u_1))$$

our pb. $\inf_{u \in U} F(u)$ is equivalent to

$(\hat{u}_1, f(\hat{u}_1))$ where $\hat{u}_1$ is ~~the~~ solves

$$\inf_{u_1 \in O_1} \tilde{F}(u_1) \quad \underbrace{\tilde{F}(u_1) = F(u_1, f(u_1))}_{\hat{u}}$$

$$\Rightarrow 0 = D\tilde{F}(\hat{u}_1) =$$

$$0 = D_1 F(\underline{\underline{\hat{u}}}) + D_2 F(\hat{u}) \, Df(\hat{u}_1)$$

$$D_1 F(\hat{u}) = D_2 F(\hat{u}) (D_2 G(\hat{u}))^{-1} D_1 G(\hat{u})$$

$$D_2 F(\hat{u}) = D_2 F(\hat{u}) (D_2 G(\hat{u}))^{-1} D_2 G(\hat{u})$$

$$DF(u) = -\lambda(\bar{u}) \qquad DG(\bar{u})$$