

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

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

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

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

(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 $(*) \Leftrightarrow \nabla F(\hat{u}) = 0$

The projection thm.

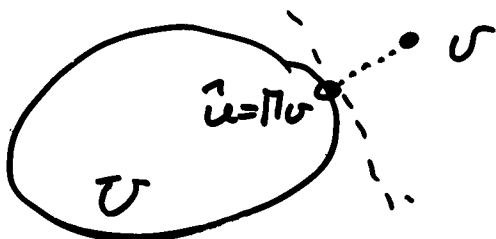
V Hilbert, real $U \subset V$ nonempty convex closed

If $v \in V$, then (i) \Leftrightarrow (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

$u \in U$
convex

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

By Euler :

$$\begin{aligned} \nabla F(u) &= u - v & (F(u) = \frac{1}{2} \langle u - v, u - v \rangle) \\ \nabla^2 F &= I \end{aligned}$$

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

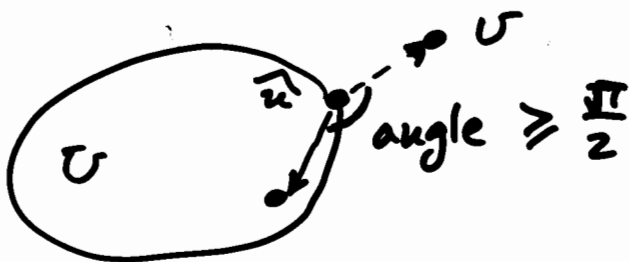
so we proved (i).

$$\|v - \hat{u}\| \leq \|v - u\| \quad \forall u \in U$$

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

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

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



Defn.

Π is called the projection operator.

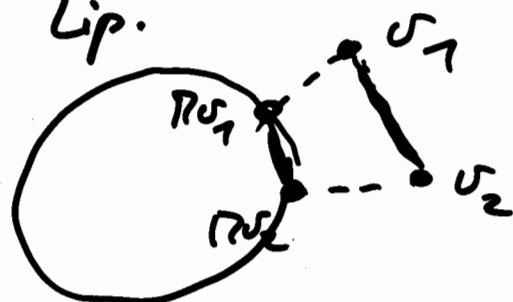
Properties of $\Pi: V \rightarrow U$

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

\rightarrow (ii) $\|\Pi v_1 - \Pi v_2\| \leq \|v_1 - v_2\| \quad \forall v_1, v_2 \in V$

$\Rightarrow \Pi$ 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)

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

$= \langle \Pi v_1 - \Pi v_2, \Pi v_1 - v_1 \rangle \leq 0$ by proj thm.

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

$+ \langle \Pi v_1 - \Pi v_2, v_2 - \Pi v_2 \rangle \leq 0$ by proj thm.

$\leq \|\Pi v_1 - \Pi v_2\| \cdot \|v_1 - v_2\|$ (Schwarz) \Rightarrow (ii)

if U vector subspace

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

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

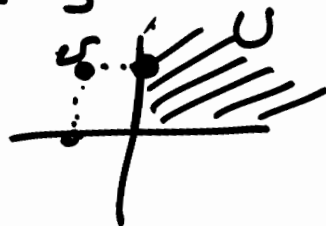
$\langle v - \hat{u}, \hat{u} - u \rangle \leq 0$

$\Rightarrow = 0.$

Examples

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

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



$$U = \{ u \in \mathbb{R}^n : a_i \leq u[i] \leq b_i \}$$

⊛

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

Ⓠ) The relaxation method can be applied \S :
when U is as above.

$$F(u_{k+1}(1), \dots, u_{k+1}(i), \dots, u_k(n)) =$$

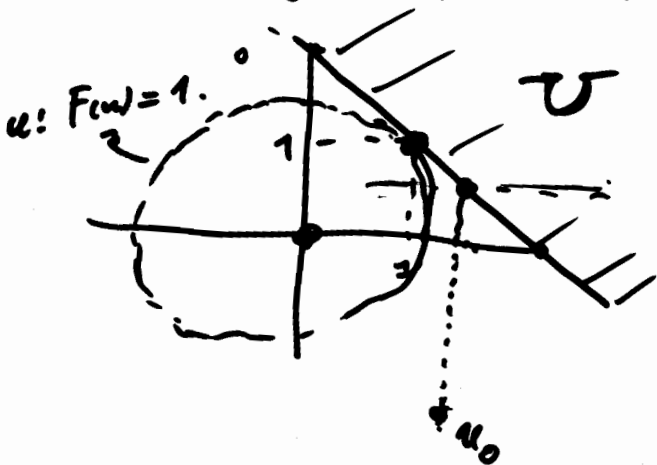
$$= \inf_{a_i \leq \rho \leq b_i} F(u_{k+1}(1), \dots, \rho, \dots, u_k(n))$$

u_k

Cannot be extended to more general constraints

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

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



Then F elliptic and U
is of the form ⊛
then u_k converges.

2.3.3. Projected gradient with a variable step.

U convex closed F convex

Motivation.

$$\hat{u} \in U \text{ and } F(\hat{u}) = \inf_{u \in U} F(u)$$

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

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

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

\hat{u} is the projection of u !

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

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

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

we will use the method of successive projections:

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

If $U = V$ then $\Pi = \text{identity}$... the grad. method.

G is a contraction if $\exists \gamma \in (0, 1)$

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

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

(V complete)

We will choose ρ such that G is a contraction.

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

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

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

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

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

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

then (u_k) converges to

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

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

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

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

We have to show that G_k is a contraction $\forall k$.
 $u_1, u_2 \in V$ arbitrary:

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

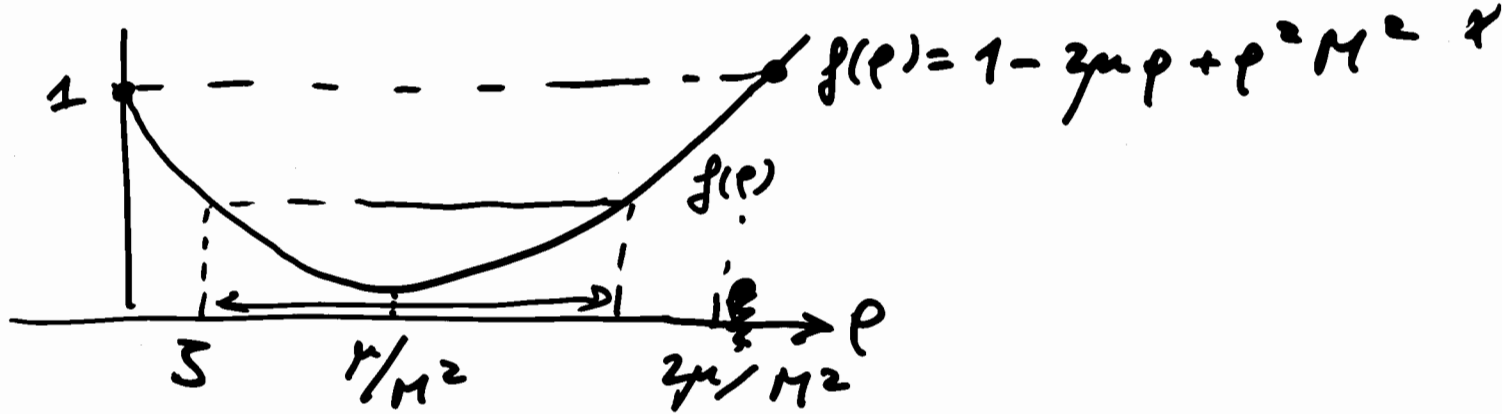
$$\leq \|u_1 - u_2 - \rho_k (\nabla F(u_1) - \nabla F(u_2))\|^2$$

$$= \|u_1 - u_2\|^2 - 2\rho_k \langle \nabla F(u_1) - \nabla F(u_2), u_1 - u_2 \rangle + \rho_k^2 \|\nabla F(u_1) - \nabla F(u_2)\|^2$$

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

$$= (1 - 2\mu\rho_k + \rho_k^2 M^2) \|u_1 - u_2\|^2$$

must be < 1
 (and > 0)



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

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

Rem. the cond. on ρ_k is sufficient.

Gradient with projection for an elliptic F.

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

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

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

$$\|\nabla F(u_1) - \nabla F(u_2)\| = \|B(u_1 - u_2)\| \geq \underbrace{\lambda_{\min}(B)}_{\mu} \|u_1 - u_2\|$$

$$\|\nabla F(u_1) - \nabla F(u_2)\| \leq \underbrace{\lambda_{\max}(B)}_M \|u_1 - u_2\|$$

The thm recommends

$$\delta \in \left(0, \frac{\lambda_{\min}}{\lambda_{\max}^2}\right) \quad \delta \leq \rho_k \leq \frac{2\lambda_{\min}}{\lambda_{\max}^2} - \delta$$

We can improve

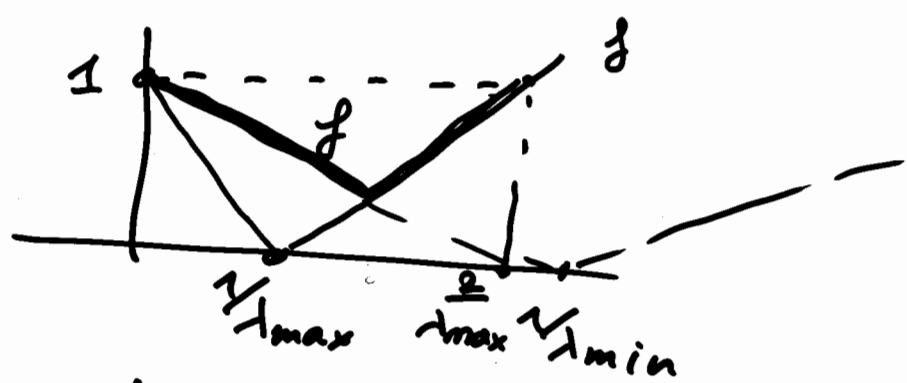
$$\begin{aligned} u_{k+1} - \hat{u} &= u_k - \hat{u} - \rho_k B(u_k - \hat{u}) \\ &= (I - \rho_k B)(u_k - \hat{u}) \end{aligned}$$

$$\|u_{k+1} - \hat{u}\| \leq \|I - \rho_k B\|_2 \|u_k - \hat{u}\|$$

symmetric

its spectral radius is the max. eigenvalue

$$f(\rho) = \max \left\{ |1 - \rho \lambda_{\min}(B)|, |1 - \rho \lambda_{\max}^8(B)| \right\}$$



→ 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) $G: \mathbb{R}^n \rightarrow \mathbb{R}$
 continuous, convex $G(u) \geq 0 \quad \forall u \in \mathbb{R}^n$
 and $G(u) = 0 \iff u \in U$

For every $\varepsilon > 0$

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

We find the solution \bar{u} by letting $\varepsilon \rightarrow 0$

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

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

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

Proof. $F_\varepsilon(u) \geq F(u) \Rightarrow F_\varepsilon$ coercive;
 strictly convex \Rightarrow (i)

$$F(u_\varepsilon) \leq \underbrace{F(u_\varepsilon) + \frac{1}{\varepsilon} G(u_\varepsilon)} = \mathbb{F}_\varepsilon(u_\varepsilon) \leq \mathbb{F}_\varepsilon(\hat{u}) = \underline{F(\hat{u})}$$

all. \mathbb{F}_ε concave... $(u_\varepsilon)_{\varepsilon > 0}$ is bounded.

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

$$u_{\varepsilon_k} \xrightarrow{k \rightarrow \infty} \tilde{u}$$

$$F(u_{\varepsilon_k}) \leq F(\hat{u}) \quad \Bigg| \quad F(\tilde{u}) = \lim_{k \rightarrow \infty} F(u_{\varepsilon_k}) \leq F(\hat{u})$$

F continuous

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

\uparrow
goes to 0

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

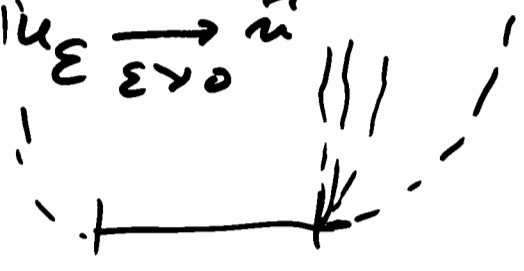
$$\Rightarrow \tilde{u} \in U$$

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

we have a unique solution

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

$$u_\varepsilon \xrightarrow{\varepsilon > 0} \hat{u}$$



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.

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

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

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

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

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

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

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

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

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

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

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

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

11

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

$$\inf_{u_1 \in Q_1} \tilde{F}(u_1) \quad \tilde{F}(u_1) = \underbrace{F(u_1, f(u_1))}_{\hat{u}}$$

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

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

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

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

$$DF(u) = -\lambda(\hat{u}) \quad DG(\hat{u})$$