Optimization

Introduction to optimization problems
- minima, maxima, feasible sets, equivalent reformulations

Criteria for optimality
- differentiable case; sufficient & necessary conditions

Duality
- Lagrange function, dual problem

Convex analysis and optimization
- convex sets, convex functions, optimality criteria

Optimization algorithms
- simplex method
- primal-dual methods
- smooth unconstrained optimization
- constrained optimization

Optimization problem

$$\min f(x), \quad x \in \mathbb{R}^n \quad \text{subject to} \quad c_E(x) = b_E \text{ and } c_I(x) \le b_I$$

x — optimisation variable

$f : \mathbb{R}^n \longrightarrow \mathbb{R}$ — objective function

$c_E : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ — equality constraint function

$c_I : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ — inequality constraint function

$b_E \in \mathbb{R}^p$

$b_I \in \mathbb{R}^m$

$c_I(x) \le b_I$ is meant component-wise

Classes of optimization problems

linear program : f, c_E, c_I linear/affine

nonlinear program : f, c_E, c_I nonlinear

convex optimisation problem : f, c_E linear, c_I convex (comp.-wise)

$g: \mathbb{R}^n \to \mathbb{R}^m$ linear $\iff g(ax + by) = a\,g(x) + b\,g(y) \;\forall x, y \in \mathbb{R}^n, a, b \in \mathbb{R}$

$g: \mathbb{R}^n \to \mathbb{R}$ convex $\iff g(ax + (1-a)y) \leq a\,g(x) + (1-a)\,g(y) \;\forall x, y \in \mathbb{R}^n, a \in [0,1]$

other types not covered in this lecture (at least initially)

discrete optimisation problem : f, c_E, c_I defined on $\mathbb{Z}^n$

combinatorial optimisation pb. : f, c_E, c_I defined on $\{1, \dots, k\}^n$

optimisation in Banach spaces : f, c_E, c_I def. on Banach sp.

optimisation on manifolds : f, c_E, c_I def. on Riemann. manif.

Applications

portfolio optimisation

- optimally invest money in n stocks; x_i = investment in stock i
- constraint on budget, positivity, ...
- objective fcn: total risk, variance of income, ...

process- or product-optimisation

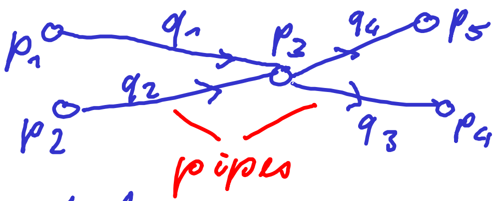
- x_i = production amount of product i
- constraints on ressources, capacity, ...
- objective fcn: profit

data fitting

- x = parameter in a model
- constraints: a-priori information (positivity etc.)
- objective fcn: data discrepancy

Example 1: Optimization of high pressure gas network

- network



q_i = flows
p_i = pressures
d_i = demands

- mass constraint for node 3: $\quad q_1 + q_2 - q_3 - q_4 - d_3 = 0$

  in general $\quad Aq - d = 0 \quad$ linear constraint

  sparse matrix

- pipe constraint for pipe 1: $\quad p_3^2 - p_1^2 + h_1 q_1^\alpha = 0$

  in general $\quad p^T B p + g(q) = 0 \quad$ quadratic / nonlinear constraint

- compressor constraint: $\quad q_6 - q_7 + z\, h(p_6, q_6, p_7, q_7) \geq 0$

  in general $\quad Cq + z \cdot h(p, q) \geq 0 \quad$ discrete / nonlinear constraint

$$z = \begin{cases} 0 & \text{if machine off} \\ 1 & \text{if on} \end{cases}$$

- bounds $\quad p \leq p_{max}$
- objectives: minimise supply / compressor costs, etc.

Example 2: Least squares fitting

$$\min \|Ax - b\|_2^2 = \sum_{i=1} (a_i^T x - b_i)^2, \quad A \in \mathbb{R}^{k \times n}, \; b \in \mathbb{R}^k, \; k \geq n$$

rows of A

seek minimum by setting derivative to zero:

$$0 = A^T(Ax - b)$$

if A^TA is invertible, then

$$x = (A^TA)^{-1} A^T b$$

pseudoinverse $A^+ = (A^TA)^{-1} A^T$

classical image processing application: deblurring

x = pixel-wise vector of grey values of true image

b = observed (blurred) grey values

A = blurring matrix

Difference to Calculus of Variations

- both fields tightly connected
- CoV tries to establish existence of minimisers and their properties, typically via "energy methods"
- Optimisation tries to find minimisers and criteria for optimality; emphasises numerical methods

Overall plan of lecture

- Introduction to optimisation problems
  (minima, maxima, feasible sets, equivalent reformulations)
- Criteria for optimality
  (differentiable case; sufficient & necessary conditions)
- Duality
  (Lagrange function, dual problem)
- Convex optimisation
  (convex sets, convex functions, optimality criteria)
- Optimisation algorithms
  - simplex method
  - primal-dual methods
  - smooth unconstrained optimisation
  - Constrained optimisation

Basic notation

$$\min \; f(x) \quad \text{subject to} \quad c_E(x) = 0, \; c_I(x) \leq 0 \;\; (\text{pt-wise}) \qquad (*)$$

$(*)$ describes the problem to find $x \in \mathbb{R}^n$ that minimises f among all x satisfying the constraints

notation:
- $f \equiv f_0 : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\infty\}$
- $c_I \equiv (f_1, \dots, f_m)^T : \mathbb{R}^n \longrightarrow (\mathbb{R} \cup \{\infty\})^m$
- $c_E \equiv (h_1, \dots, h_p)^T : \mathbb{R}^n \longrightarrow (\mathbb{R} \cup \{\infty\})^p$

<u>Def:</u> <u>domain of f</u> $: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$: $\operatorname{dom} f = \{x \in \mathbb{R}^n \mid f(x) < \infty\}$

<u>Def:</u> <u>domain of optimisation problem</u> $(*)$: $\mathcal{D} = \bigcap_{i=0}^{m} \operatorname{dom} f_i \cap \bigcap_{i=1}^{p} \operatorname{dom} h_i$

Basic terminology

Def: • $x \in \mathcal{D}$ is feasible if all constraints $f_1(x), \dots, f_m(x) \le 0$ and $h_1(x) = \dots = h_p(x) = 0$ are fulfilled.

• $(*)$ is feasible if there is a feasible $x \in \mathcal{D}$

Def: • $p^* = \inf \{ f_0(x) \mid x \in \mathcal{D} \text{ is feasible} \}$ is called optimal value

• if $(*)$ is not feasible, we set $p^* := \infty$

• if $\exists\, x_k \in \mathcal{D}$ feasible with $f_0(x_k) \xrightarrow[k \to \infty]{} -\infty$, set $p^* = -\infty$; then $(*)$ is called unbounded from below

Def: • x^* is called (globally) optimal point if x^* is feasible and $f(x^*) = p^*$

• $X_{opt} = \{ x \in \mathcal{D} \mid x \text{ is optimal point} \}$

• A feasible point x with $f_0(x) \le p^* + \varepsilon$ is called ε-suboptimal

• A feasible point x is called locally optimal if $\exists\, R > 0$ s.t.

$$f_0(x) = \inf \{ f_0(z) \mid z \in \mathcal{D} \text{ is feasible},\ \|z - x\| \le R \}$$

• A locally/globally optimal point x is called strictly optimal if $\exists\, R > 0$ s.t. $f_0(x) < f_0(y)$ for all feasible $y \in \mathcal{D}$ with $y \ne x$, $\|x - y\| \le R$

Examples

Optimisation problems on $\mathbb{R}_{++} := \{ x \in \mathbb{R} \mid x > 0 \}$

1) $f_0(x) = \frac{1}{x}$ $\quad p^* = 0 \,, \; X_{opt} = \emptyset$

2) $f_0(x) = -\log x$ $\quad p^* = -\infty \,, \; X_{opt} = \emptyset$

3) $f_0(x) = x \log x$ $\quad p^* = -\frac{1}{e} \,, \; x^* = \frac{1}{e} \,, \; X_{opt} = \{ \frac{1}{e} \}$

Thm: Let $(*)$ be a convex optimisation problem, i.e. $f_0, \dots, f_m$ are convex, $h_1, \dots, h_p$ affine. Then x^* locally optimal $\Longleftrightarrow$ x^* globally optimal

proof: homework

Equivalent optimization problems

- $(*)$ is the standard form of an optimisation problem.
  It can always be generated, e.g. the constraint $g_i(x) = \tilde{g}_i(x)$ is equivalent to $h_i(x) := g_i(x) - \tilde{g}_i(x) = 0$ or $\tilde{f}_i(x) \geq 0$ is equivalent to $f_i(x) := -\tilde{f}_i(x) \leq 0$.

- Two optimisation problems are called equivalent if the solution of one can be calculated from the solution of the other

Example: $\min \tilde{f}_0(x) = \alpha_0 f_0(x)$

subject to $\tilde{f}_i(x) = \alpha_i f_i(x) \leq 0, \tilde{h}_j(x) = \beta_j h_j(x) = 0, i = 1, \dots, m, j = 1, \dots, p$

is for $\alpha_i > 0$, $\beta_j \neq 0$ equivalent to

$\min f_0(x)$ subject to $f_i(x) \leq 0, h_j(x) = 0$

Basic transformations I

(P_1) $\min f_0(x)$ s.t. $f_i(x) \leq 0$, $i = 1, \dots, m$, $h_i(x) = 0$, $i = 1, \dots, p$

(P_2) $\min \tilde{f}_0(x)$ s.t. $\tilde{f}_i(x) \leq 0$, $i = 1, \dots, m$, $\tilde{h}_i(x) = 0$, $i = 1, \dots, p$

- Substitution of variables

Let $\phi : \mathbb{R}^n \to \mathbb{R}^n$ bijection,

$\tilde{f}_i = f_i \circ \phi$, $i = 0, \dots, m$, $\tilde{h}_i = h_i \circ \phi$, $i = 1, \dots, p$

If x solves (P_1), then $\phi^{-1}(x)$ solves (P_2) and vice versa.

- Transformation of objective and constraint functions

Let $\psi_0 : \mathbb{R} \to \mathbb{R}$ strictly monotonically increasing

$\psi_1, \dots, \psi_m : \mathbb{R} \to \mathbb{R}$ with $\psi_i(u) \leq 0 \iff u \leq 0$

$\phi_1, \dots, \phi_p : \mathbb{R} \to \mathbb{R}$ with $\phi_i(u) = 0 \iff u = 0$

$\tilde{f}_i = \psi_i \circ f_i$, $i = 0, \dots, m$, $\tilde{h}_i = \phi_i \circ h_i$, $i = 1, \dots, p$

We have $X_{opt}^{(P_1)} = X_{opt}^{(P_2)}$

Basic transformations II

(P_1) $\min f_0(x)$ s.t. $f_i(x) \leq 0$, $i=1,\dots,m$, $h_i(x)=0$, $i=1,\dots,p$

- slack variables s_i: exploit $f_i(x) \leq 0 \iff \exists\, s_i \geq 0 : f_i(x)+s_i=0$

(P_2) $\min f_0(x)$ s.t. $s_i \geq 0$, $i=1,\dots,m$, $f_i(x)+s_i=0$, $h_j(x)=0$, $i=1,\dots,m$, $j=1,\dots,p$

(x,s) feasible (optimal) for $(P_2) \iff x$ feasible (optimal) for (P_1) & $s_i := -f_i(x)$

- Elimination of equality constraints

Assume $C=\{x \in \mathbb{R}^n \mid h_1(x)=\dots=h_p(x)=0\}$ can be parametrised by $\phi : \mathbb{R}^k \to \mathbb{R}^n$, i.e. $x \in C \iff \exists\, z \in \mathbb{R}^k : x=\phi(z)$

$\tilde{f}_i = f_i \circ \phi$, $i=0,\dots,m$

(P_2) $\min \tilde{f}_0(x)$ s.t. $\tilde{f}_1(x),\dots,\tilde{f}_m(x) \leq 0$

$$X_{opt}^{(P_1)} = \phi\left(X_{opt}^{(P_2)}\right)$$

Example: $h_1(x)=\dots=h_p(x)=0 \iff Ax=b$ (linear constraints)

Set $\phi(z) = Fz + x_0$ for $F \in \mathbb{R}^{n\times k}$ with range $F = \ker A$, x_0 feasible

Basic transformations III

(P_1) $\min f_0(x)$ s. t. $f_i(x) \leq 0$, $i=1,\dots,m$, $h_i(x)=0$, $i=1,\dots,p$

- implicit constraints

Let $F(x) = \begin{cases} f_0(x) & \text{if } f_1(x),\dots,f_m(x) \leq 0,\ h_1(x)=\dots=h_p(x)=0 \\ \infty & \text{else} \end{cases}$

(P_2) $\min F(x)$

$X_{opt}^{(P_2)} = X_{opt}^{(P_1)}$

<u>Example</u>: $\min \begin{cases} \|x\|^2 & \text{if } Ax=b \\ \infty & \text{else} \end{cases} \iff \min \|x\|^2$ s. t. $Ax=b$

- optimisation over one variable

Let $x \equiv (x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$; $\tilde{F}(x_1) = \inf \{ f_0(x_1,x_2) \mid f_1(x_1,x_2),\dots,f_m(x_1,x_2) \leq 0, h_1(x_1,x_2) = \dots h_p(x_1,x_2)=0 \}$

(P_2) $\min \tilde{F}(x_1)$

$x_1 \in X_{opt}^{(P_2)} \iff \exists\, x_2 \in \mathbb{R}^{n_2} : (x_1, x_2) \in X_{opt}^{(P_1)}$

useful if minimisation over x_2 can be done explicitly

Basic transformations IV

(P_1) $\min f_0(x)$ s.t. $f_i(x) \leq 0,\ i = 1, \dots, m,\quad h_i(x) = 0,\ i = 1, \dots, p$

- epigraph formulation

(P_2) $\min_{t,x} t$ s.t. $f_0(x) - t \leq 0,\ f_1(x), \dots, f_m(x) \leq 0,\ h_1(x) = \dots = h_p(x) = 0$

$(x, t) \in X_{opt}^{(P_2)} \iff x \in X_{opt}^{(P_1)}$ and $f_0(x) = t$

- generalised epigraph formulation

Let $f_0(x) = \sum_{j=1}^{M} g_j(x)$

(P_2) $\min_{t_1, \dots, t_M, x} \sum_{j=1}^{M} t_j$ s.t. $g_j(x) - t_j \leq 0,\ j = 1, \dots, M$ + other constraints

$(x, t_1, \dots, t_M) \in X_{opt}^{(P_2)} \iff x \in X_{opt}^{(P_1)}$ and $g_j(x) = t_j,\ j = 1, \dots, M$

Motivation and notation

Optimality Conditions

- indicate when a point is not optimal (necessary conditions)
- guarantee that a candidate solution is indeed (locally) optimal (sufficient condition)
- guide the design of algorithms

Assumption and notation

- in the following, we assume that $f_0, \ldots, f_m, h_1, \ldots, h_p$ are continuously differentiable up to the order needed in each result (e.g., for first order optimality conditions up to first order)
- $Df = \left(\partial f / \partial x_1, \ldots, \partial f / \partial x_n \right)$, $\nabla f = Df^T$
- $D^2 f = \begin{pmatrix} \partial^2 f / \partial x_1^2 & \cdots & \partial^2 f / \partial x_1 \partial x_n \\ \vdots & & \vdots \\ \partial^2 f / \partial x_n \partial x_1 & \cdots & \partial^2 f / \partial x_n^2 \end{pmatrix}$

Optimality criteria: Unconstrained optimization case $m = p = 0$

1st order necessary conditions

Thm: If x^* is a local optimum of (✱), then $Df_0(x^*) = 0$.

proof: By contradiction: suppose $Df_0(x^*) \neq 0$.

By Taylor's thm, for $\alpha > 0$,

"little o-notation"

$$f_0(x^* - \alpha \nabla f(x^*)) = f_0(x^*) + Df_0(x^*)(-\alpha \nabla f_0(x^*)) + o(\|\alpha \nabla f_0(x^*)\|)$$

$$= f_0(x^*) - \underbrace{\alpha \|\nabla f_0(x^*)\|^2}_{A} + \underbrace{o(\alpha \|\nabla f_0(x^*)\|)}_{B}$$

For all $\alpha > 0$ small enough, $B < A$ so that

$$f_0(x^* - \alpha \nabla f(x^*)) < f_0(x^*)$$

$\Rightarrow$ x^* cannot be local optimum. $\square$

"little-o notation": We write $o(g(s))$ if we want to refer to a function $f(s)$ with $\lim_{s \to 0} \frac{f(s)}{g(s)} = 0$.

Optimality criteria: Unconstrained optimization

2nd order necessary conditions

Thm: If x^* is a local optimum of (*), then $Df_0(x^*) = 0$ and $D^2 f_0(x^*)$ is positive semi-definite, i.e.

$$s^T D^2 f_0(x^*) s \geq 0 \quad \forall s \in \mathbb{R}^n.$$

proof: By contradiction: suppose $\exists s \in \mathbb{R}^n : s^T D^2 f_0(x^*) s < 0$.

We know $Df_0(x^*) = 0$ already.

By Taylor's thm, for $\alpha > 0$,

$$f_0(x^* + \alpha s) = f_0(x^*) + \underbrace{\frac{1}{2}(\alpha s)^T D^2 f_0(x^*)(\alpha s)}_{A < 0} + \underbrace{o(|\alpha s|^2)}_{B}$$

For all $\alpha > 0$ small enough, $B < |A|$ so that

$$f_0(x^* + \alpha s) < f_0(x^*)$$

$\Rightarrow$ x^* cannot be local optimum. $\square$

2nd order sufficient conditions

Thm: If $Df_0(x^*) = 0$ and $D^2 f_0(x^*)$ is positive definite, i.e.

$$s^T D^2 f_0(x^*) s > 0 \qquad \forall s \in \mathbb{R}^n,$$

then x^* is a strictly locally optimal point.

proof: By continuity, $D^2 f_0(x)$ is pos. def. for all x in an open ball $B_\delta(x^*)$ around x^*.

Let $x \in B_\delta(x^*)$, $x \neq x^*$. By Taylor's thm., $\exists\, z$ between x and x^* s.t.

$$f_0(x) = f_0(x^*) + \underbrace{Df_0(x^*)(x - x^*)}_{0} + \underbrace{\tfrac{1}{2}(x-x^*)^T D^2 f_0(z)(x - x^*)}_{>0}$$

$$> f_0(x^*)$$

$\square$

Example in R^2

Find local minima of $f(x_1, x_2) = x_1^3 + 3x_1^2 x_2 + 3x_1 x_2^2 + x_2^3 - 2x_1^2 - 8x_1 x_2 - 2x_2^2$

1st order necessary cond.: $0 = \nabla f(x_1, x_2) = \begin{pmatrix} 3x_1^2 + 6x_1x_2 + 3x_2^2 - 4x_1 - 8x_2 \\ 3x_2^2 + 6x_1x_2 + 3x_1^2 - 4x_2 - 8x_1 \end{pmatrix}$

$\Longrightarrow x = (0,0)$ or $x = (1,1)$

2nd order necessary cond: $D^2 f(x_1, x_2) = \begin{pmatrix} 6x_1 + 6x_2 - 4 & 6x_1 + 6x_2 - 8 \\ 6x_1 + 6x_2 - 8 & 6x_1 + 6x_2 - 4 \end{pmatrix}$

local maximum or saddle?

$D^2 f(0,0) = \begin{pmatrix} -4 & -8 \\ -8 & -4 \end{pmatrix}$ not pos. semi-def. (determinant < 0)

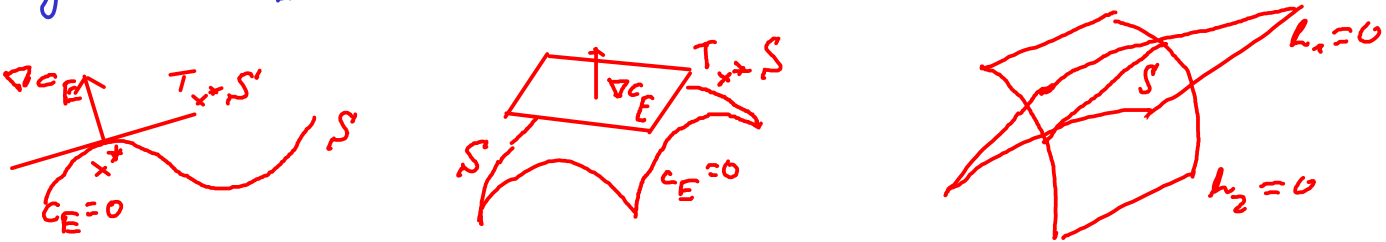
2nd order sufficient cond: $D^2 f(1,1) = \begin{pmatrix} 8 & 4 \\ 4 & 8 \end{pmatrix}$ pos def. (all minors pos.)

$\Longrightarrow (1,1)$ is only locally optimal point

Tangent plane & regular points

A set of equality constraints $0 = c_E(x) = \begin{pmatrix} h_1(x) \\ \vdots \\ h_p(x) \end{pmatrix}$ defines a hypersurface S in $\mathbb{R}^n$.

In general: c_E smooth $\Rightarrow$ S smooth ← what can happen at non-regular points?



Def: • A <u>curve on S</u> is a continuous function $x : [a,b] \to S$.

• x "<u>passes through</u> $x^* \in S$" if $\exists\, t^* \in [a,b] : x(t^*) = x^*$.

• The <u>tangent plane</u> $T_{x^*}S$ to S at x^* is the vector space

$$T_{x^*}S = \{ \dot{x}(t^*) \mid x \text{ is differentiable curve on } S \text{ with } x(t^*) = x^* \}$$

Def: $x^* \in \mathbb{R}^n$ with $c_E(x^*) = 0$ is called <u>regular</u> if $\nabla h_1(x^*), \dots, \nabla h_p(x^*)$ are linearly independent.

Tangent plane at regular points

Thm: At a regular point x^* of the surface $S = \{x \in \mathbb{R}^n \mid c_E(x) = 0\}$,

$$T_{x^*}S = \{y \mid Dc_E(x^*)\, y = 0\}.$$

proof: "$\subset$": Let $y \in T_{x^*}S$, i.e. $\exists\, x : [a,b] \to S$, $t^* \in [a,b]$, $x(t^*) = x^*$, $\dot{x}(t^*) = y$.

$$c_E(x(t)) = 0 \implies 0 = [c_E(x(t))]'\Big|_{t=t^*} = Dc_E(x(t^*))\,\dot{x}(t^*)$$

"$\supset$": Let $Dc_E(x^*)\,y = 0$.

idea: find a curve $t \mapsto x^* + ty + Dc_E(x^*)^T u(t)$ on S

Consider $F : \mathbb{R}^{p+1} \to \mathbb{R}^p$, $\quad F(t,u) = c_E(x^* + ty + Dc_E(x^*)^T u)$

$F(0,0) = 0$, $\quad D_u F(0,0) = Dc_E(x^*)\, Dc_E(x^*)^T$ is invertible

$\xRightarrow{\text{implicit fcn thm}}$ $\exists$ continuous $u : [-a,a] \to \mathbb{R}^p$ s.t. $F(t,u(t)) = 0$

$\implies \quad x(t) = x^* + ty + Dc_E(x^*)^T u(t)$ is curve on S

$$\implies 0 = \frac{d}{dt} c_E(x(t))\Big|_{t=0} = Dc_E(x^*)\,y + Dc_E(x^*)\,Dc_E(x^*)^T \dot{u}(0)$$

$$\implies \dot{u}(0) = 0 \implies \dot{x}(0) = y + Dc_E(x^*)^T \dot{u}(0) = y \qquad \square$$

Case $m=0$

1st order necessary conditions I

Lemma: If x^* is a local optimum of (*) and regular w.r.t. c_E, then $Df_0(x^*)\,y = 0$ for all $y \in \mathbb{R}^n$ with $Dc_E(x^*)\,y = 0$.

proof: • Let $Dc_E(x^*)\,y = 0 \xRightarrow{\text{previous thm}} y \in T_{x^*}S$ for $S = \{x \in \mathbb{R}^n \mid c_E(x) = 0\}$

$\Rightarrow \exists\, x : [-a, a] \to S$ differentiable, $\dot{x}(0) = y$, $x(0) = x^*$.

• x^* is local optimum of (*)

$\Rightarrow t = 0$ is local minimizer of $t \mapsto f_0(x(t))$

$\Rightarrow 0 = [f_0(x(t))]'\big|_{t=0} = Df_0(x^*)\,y$ $\square$

interpretation: $\nabla f_0(x^*)$ is orthogonal to tangent plane

1st order necessary conditions II

Thm: If x^* is a local optimum of (*) and regular w.r.t. c_E, then $\exists$ Lagrange multiplier $\lambda \in \mathbb{R}^p$ s.t. $Df_0(x^*) + \lambda^T Dc_E(x^*) = 0$.

proof: Abbreviate $A = Dc_E(x^*)$ and $g = \nabla f_0(x^*)$

Let the columns of N be a basis of $\ker A = (\text{range}\, A^T)^\perp$

$\Rightarrow \quad g = -A^T\lambda + Nz$ for some $\lambda \in \mathbb{R}^p$, $z \in \mathbb{R}^{n-p}$

The previous Lemma implies $N^T g = 0$, i.e.

$$0 = -N^T A^T \lambda + N^T N z = N^T N z.$$

However, N has full rank so that $z = 0$.

$$\Rightarrow \quad g = -A^T \lambda \qquad \square$$

interpretation: $\nabla f_0(x^*)$ is lin. comb. of $\nabla h_1(x^*), \dots, \nabla h_p(x^*)$

2nd order necessary conditions

Thm: If x^* is a local optimum of $(*)$ and regular w.r.t. c_E, then $\exists \lambda \in \mathbb{R}^p$ with $Df_0(x^*) + \lambda^T Dc_E(x^*) = 0$ and

$$D^2 f_0(x^*) + \sum_{k=1}^{p} \lambda_k D^2 h_k(x^*)$$

is positive semidefinite on $T_{x^*}S$ for $S = \{x \in \mathbb{R}^n \mid c_E(x) = 0\}$

proof: Let $x(t)$ be a curve on S with $x(0) = x^*$.

By 2nd order condition for unconstrained problems,

$$0 \le \frac{d^2}{dt^2} f_0(x(t))\Big|_{t=0} = \dot{x}(0)^T D^2 f_0(x(0)) \dot{x}(0) + Df_0(x(0))\, \ddot{x}(0).$$

Also, $0 = \lambda^T c_E(x(t))$ implies

$$0 = \frac{d^2}{dt^2} \lambda^T c_E(x(t)) = \sum_{k=1}^{p} \lambda_k \dot{x}(0)^T D^2 h_k(x(0)) \dot{x}(0) + \lambda^T Dc_E(x(0))\, \ddot{x}(0)$$

Adding both eqs. and using the 1st order condition,

$$0 \le \dot{x}(0)^T \Big[D^2 f_0(x^*) + \sum_{k=1}^{p} \lambda_k D^2 h_k(x^*) \Big] \dot{x}(0),$$

where $\dot{x}(0) \in T_{x^*}S$ was arbitrary. $\square$

2nd order sufficient conditions

Thm: If there are $x^* \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^p$ with

$$c_E(x^*) = 0, \quad Df_0(x^*) + \lambda^T Dc_E(x^*) = 0,$$

$$D^2 f_0(x^*) + \sum_{k=1}^{p} \lambda_k D^2 h_k(x^*) \text{ pos. def. on } T_{x^*}S,$$

then x^* is a strict local optimum of $(*)$.

proof: Assume the opposite, i.e. $\exists$ sequence $y_k = x^* + \delta_k s_k$, $\delta_k > 0$, $\delta_k \xrightarrow[k\to\infty]{} 0$, $|s_k| = 1$, with $c_E(y_k) = 0$ and $f_0(y_k) \leq f_0(x^*)$.

s_k bounded $\Longrightarrow$ subsequence converges. Wlog, $s_k \longrightarrow s^*$.

$0 = c_E(y_k) - c_E(x^*)$; divide this by δ_k and let $k \to \infty$ $\Longrightarrow Dc_E(x^*)s^* = 0$.

Taylor's thm: $0 \geq f_0(y_k) - f_0(x^*) = \delta_k Df_0(x^*)s_k + \frac{\delta_k^2}{2} s_k^T D^2 f_0(\eta_0) s_k \quad (E_0)$

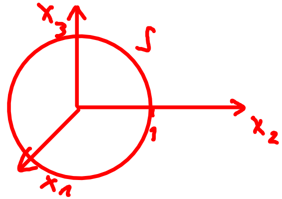
$0 = h_i(y_k) - h_i(x^*) = \delta_k Dh_i(x^*)s_k + \frac{\delta_k^2}{2} s_k^T D^2 h_i(\eta_i) s_k \quad (E_i)$

for some η_i between x^* and y_k.

$(E_0) + \sum_{i=1}^{p} \lambda_i (E_i) \Longrightarrow 0 \geq \frac{\delta_k^2}{2} s_k^T \left[D^2 f_0(\eta_0) + \sum_{i=1}^{p} \lambda_i D^2 h_i(\eta_i) \right] s_k \quad \forall k \quad \lightning \quad \square$

Example

Solve $\min\limits_{x\in\mathbb{R}^3} |x-(0,2,0)|^2$ s. t. $h_1(x)=|x|^2-1=0$, $h_2(x)=x_1=0$



1st order necessary condition: $0=\nabla f(x)+\lambda_1\nabla h_1(x)+\lambda_2\nabla h_2(x) = 2\left(x-\begin{pmatrix}0\\2\\0\end{pmatrix}\right)+2\lambda_1 x+\lambda_2\begin{pmatrix}1\\0\\0\end{pmatrix}$

$0=h_1(x)$

$0=h_2(x)$

$\Longrightarrow \quad x_1=0,\quad x_3=0,\quad x_2=\pm 1,\quad \lambda_2=0,\quad \lambda_1=\begin{cases}1\\-3\end{cases}$

2nd order necessary condition: $D^2 f\begin{pmatrix}0\\-1\\0\end{pmatrix}+\lambda_1 D^2 h_1\begin{pmatrix}0\\-1\\0\end{pmatrix}+\lambda_2 D^2 h_2\begin{pmatrix}0\\-1\\0\end{pmatrix} = 2I+2\lambda_1 I \ll 0$

$\Longrightarrow$ no local minimum

2nd order sufficient condition: $D^2 f\begin{pmatrix}0\\1\\0\end{pmatrix}+\lambda_1 D^2 h_1\begin{pmatrix}0\\1\\0\end{pmatrix}+\lambda_2 D^2 h_2\begin{pmatrix}0\\1\\0\end{pmatrix} = 2I+2\lambda_1 I \gg 0$

$\Longrightarrow$ local minimum

now $m > 0$, $p > 0$

Active set & regular points

Def: • Let x^* be feasible, i.e. $h_1(x^*) = \ldots = h_p(x^*) = 0$, $f_1(x^*), \ldots, f_m(x^*) \le 0$.

Let $\mathcal{J}$ be the set of indices j with $f_j(x^*) = 0$.

The set of constraints $f_j(x^*) \le 0$, $j \in \mathcal{J}$, is called active set.

$f_j(x^*) \le 0$, $j \notin \mathcal{J}$, is called inactive set.

Sometimes we count the equality contraints to the active set.

• x^* is called regular point if $\nabla h_1(x^*), \ldots, \nabla h_p(x^*)$, $\nabla f_j(x^*)$, $j \in \mathcal{J}$, are linearly independent.

In a neighbourhood of x^*, the inactive set may be completely ignored!

Optimality criteria: Inequality constraints

Karush-Kuhn-Tucker (KKT) conditions

Thm: If x^* is a locally optimal point of (*) and regular, then there are $\lambda \in \mathbb{R}^p$, $\mu \in \mathbb{R}^m$ with

$$c_E(x^*) = 0, \quad c_I(x^*) \le 0, \quad \mu \ge 0, \quad \mu^T c_I(x^*) = 0 \text{ (comp.-wise)}$$

$$Df_0(x^*) + \lambda^T Dc_E(x^*) + \mu^T Dc_I(x^*) = 0$$

"complementary slackness"

i.e. at least one of μ_i, $(c_I(x^*))_i$ is zero!

proof: • Complementary slackness

$\iff \mu_i \neq 0$ only if $i \in J$ for the active set J

• x^* is locally optimal

$\Rightarrow$ it is also locally optimal for the constraints $c_E(x) = 0$, $f_i(x) = 0$, $i \in J$

$\Rightarrow 0 = Df_0(x^*) + \lambda^T Dc_E(x^*) + \mu^T Dc_I(x^*)$ for some $\lambda \in \mathbb{R}^p$, $\mu \in \mathbb{R}^m$ with $\mu_i = 0$ if $i \notin J$

• assume $\mu_k < 0$; let $S = \{x \in \mathbb{R}^n \mid c_E(x) = 0,\ f_i(x) = 0\ \forall i \in J \setminus \{k\}\}$;

by regularity, $\exists y \in T_{x^*}S : Df_k(x^*)\, y < 0$; let x be curve on S with $x(0) = x^*$, $\dot{x}(0) = y$

$\Rightarrow x(t)$ feasible for $t \ge 0$ small and $\frac{d}{dt} f_0(x(t))\big|_{t=0} = Df_0(x^*)\, y = -\mu_k Df_k(x^*)\, y < 0$ ↯ $\square$

2nd order necessary conditions

Thm: If x^* is a locally optimal point of (*) and regular, then $\exists \lambda \in \mathbb{R}^p, \mu \in \mathbb{R}^m$ such that in addition to the KKT conditions

$$D^2 f_0(x^*) + \sum_{k=1}^{p} \lambda_k D^2 h_k(x^*) + \sum_{k=1}^{m} \mu_k D^2 f_k(x^*)$$

is pos. semi-definite on the tangent space to the active constraints.

proof: • x^* is locally optimal

$\Rightarrow$ it is also locally optimal for the constraints $c_E(x)=0$, $f_i(x)=0$, $i \in \mathcal{J}$

$\Rightarrow$ use 2^{nd} order necessary condition for equality constraints $\square$

2nd order sufficient conditions

Thm: x^* is a strictly locally optimal point of $(*)$ if $\exists \lambda \in \mathbb{R}^p, \mu \in \mathbb{R}^m$ s.t.

$$c_E(x^*) = 0, \quad c_I(x^*) \le 0, \quad \mu \ge 0, \quad \mu\, c_I(x^*) = 0 \text{ (comp.-wise)}$$

$$Df_0(x^*) + \lambda^T Dc_E(x^*) + \mu^T Dc_I(x^*) = 0$$

and $\quad D^2 f_0(x^*) + \sum_{k=1}^{p} \lambda_k D^2 h_k(x^*) + \sum_{k=1}^{m} \mu_k D^2 f_k(x^*)$

is positive definite on $\{y \mid Dc_E(x^*)y = 0,\ Df_i(x^*)y = 0 \text{ for all } i \text{ with } \mu_i > 0\}$.

proof: Assume x^* is not a strict optimum, but the conditions hold;

let $y_k = x^* + \delta_k s_k$ with $\delta_k > 0$, $|s_k| = 1$, $y_k \xrightarrow[k\to\infty]{} x^*$ such that $f_0(y_k) \le f_0(x^*)$

- w.l.o.g., $\delta_k \to 0$, $s_k \to s^*$, $0 \ge Df_0(x^*)s^*$, $0 = Dc_E(x^*)s^*$ ← as in case with equality constraints
- for each (active) constraint f_j with $\mu_j > 0$, $f_j(y_k) - f_j(x^*) \le 0 \implies Df_j(x^*)s^* \le 0$

→ if $Df_j(x^*)s^* < 0$ for a j, then $0 \ge Df_0(x^*)s^* = -\lambda^T Dc_E(x^*)s^* - \mu^T Dc_I(x^*)s^* > 0$ ↯

→ if $Df_j(x^*)s^* = 0$ for all active constraints, use proof for equality constraints

$\square$

Example

Solve $\min_{x \in \mathbb{R}^3} |x - (0,2,0)|^2$ s.t. $f_1(x) = |x|^2 - 1 \le 0$, $f_2(x) = x_1 - \frac{1}{2} \le 0$

1st order necessary condition: $0 = \nabla f(x) + \mu_1 \nabla f_1(x) + \mu_2 \nabla f_2(x) = 2\left(x - \begin{pmatrix}0\\2\\0\end{pmatrix}\right) + 2\mu_1 x + \mu_2 \begin{pmatrix}1\\0\\0\end{pmatrix}$

$0 \ge f_1(x)$, $0 \le \mu_1$, $\mu_1 f_1(x) = 0$

$0 \ge f_2(x)$, $0 \le \mu_2$, $\mu_2 f_2(x) = 0$

$\Longrightarrow$ $x_1 = 0$, $x_3 = 0$, $x_2 = 1$, $\mu_2 = 0$, $\mu_1 = 1$

2nd order sufficient condition: $D^2 f\begin{pmatrix}0\\1\\0\end{pmatrix} + \mu_1 D^2 f_1\begin{pmatrix}0\\1\\0\end{pmatrix} + \mu_2 D^2 f_2\begin{pmatrix}0\\1\\0\end{pmatrix} = 2I + 2\mu_1 I \gg 0$

$\Longrightarrow$ local minimum

Lagrange dual

Recall $(*)$ $\min_{x \in \mathbb{R}^n} f_0(x)$ s.t. $c_I(x) \equiv \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} \leq 0$, $c_E(x) \equiv \begin{pmatrix} h_1(x) \\ \vdots \\ h_p(x) \end{pmatrix} = 0$

with domain $\mathcal{D} = \bigcap_{i=0}^{m} \operatorname{dom} f_i \cap \bigcap_{j=1}^{p} \operatorname{dom} h_j$

Def: • The Lagrange function to $(*)$ is $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$, $\operatorname{dom} L = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$,

$$L(x, \mu, \lambda) = f_0(x) + \mu^T c_I(x) + \lambda^T c_E(x).$$

• μ, λ are called dual variables or Lagrange multiplicators of $(*)$

• x is called primal variable

• $g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$, $g(\mu, \lambda) = \inf_{x \in \mathcal{D}} L(x, \mu, \lambda)$, is called the Lagrange-dual or dual function

Thm: The dual function g is concave (i.e. $-g$ is convex).

proof: $-g$ is pointwise supremum of affine functions $\Longrightarrow$ convex $\square$

see next chapter!

Lower bound on optimal value

Thm: Let p^* be the optimal value of $(*)$. For every $\mu \ge 0, \lambda$ we have $g(\mu,\lambda) \le p^*$.

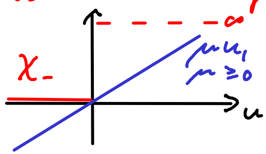
proof: Let x be feasible, i.e. $c_I(x) \le 0$, $c_E(x) = 0$, and let $\mu \ge 0$.

$$\Rightarrow g(\mu,\lambda) = \inf_{\tilde{x} \in \mathcal{D}} L(\tilde{x},\mu,\lambda) \le L(x,\mu,\lambda) = f_0(x) + \underbrace{\mu^T c_I(x) + \lambda^T c_E(x)}_{\le 0} \le f_0(x).$$

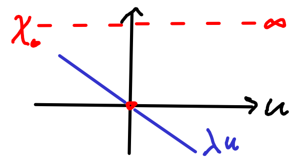
Since x was arbitrary, $g(\mu,\lambda) \le \inf_{x \text{ feasible}} f_0(x)$. □

Interpretation of dual function as "approximation":

Set $\chi_-(u) = \begin{cases} 0, & u \le 0 \\ \infty, & u > 0 \end{cases}$



$\chi_0(u) = \begin{cases} 0, & u = 0 \\ \infty, & u \ne 0 \end{cases}$



We have $(*) \iff \min_{x \in \mathbb{R}^n} \tilde{L}(x) = \min_{x \in \mathbb{R}^n} f_0(x) + \sum_{i=1}^{m} \chi_-(f_i(x)) + \sum_{j=1}^{p} \chi_0(h_j(x))$

Replace $\chi_-(u_i) \rightsquigarrow \mu_i u_i$ with $\mu_i \ge 0$, $\chi_0(u_i) \rightsquigarrow \lambda_i u_i$

(replace hard by soft constraints)

$$\Rightarrow p^* = \inf_{x \in \mathbb{R}^n} \tilde{L}(x) \ge \inf_{x \in \mathbb{R}^n} L(x,\mu,\lambda) = g(\mu,\lambda).$$

Example: least squares solution & LP

Ex: $\min_x |x|^2$ s.t. $Ax = b$

Lagrange fcn: $L(x,\lambda) = x^T x + \lambda^T(Ax - b)$ (convex, quadratic in x)

dual fcn: $g(\lambda) = \inf_{x\in\mathcal{D}} L(x,\lambda) = -\frac{1}{4}\lambda^T A A^T \lambda - \lambda^T b$ (concave, quadratic in λ)

$0 = D_x L = 2x^T + \lambda^T A \Rightarrow x = -\frac{1}{2}A^T\lambda$

property of dual fcn implies $g(\lambda) = -\frac{1}{4}\lambda^T A A^T \lambda - \lambda^T b \leq \inf_{x\in\mathcal{D}} \{x^T x \mid Ax = b\}\ \forall\lambda$

if AA^T invertible

unconstrained!

$\Rightarrow b^T(AA^T)^{-1} b = g(-2(AA^T)^{-1}b) = \max_\lambda g(\lambda) \leq \inf_{x\in\mathcal{D}}\{x^T x \mid Ax = b\}$

Ex: $\min_x c^T x$ s.t. $Ax = b$, $-x \leq 0$ $(\ast\ast)$

$L(x,\mu,\lambda) = c^T x - \mu^T x + \lambda^T(Ax - b) = -\lambda^T b + (c + A^T\lambda - \mu)^T x$

$g(\mu,\lambda) = \inf_{x\in\mathcal{D}} L(x,\mu,\lambda) = \begin{cases} -\infty & \text{if } \mu \neq c + A^T\lambda \\ -\lambda^T b & \text{else} \end{cases}$

$\Rightarrow (-\lambda^T b)$ for any λ with $\mu := c + A^T\lambda \geq 0$ is lower bound for $(\ast\ast)$.

Dual problem

Recall: Every (μ, λ) with $\mu \geq 0$ satisfies $g(\mu,\lambda) \leq p^*$.

Def: · The dual problem or Lagrange-dual problem is

$$\underbrace{\max_{\mu \in \mathbb{R}^m, \lambda \in \mathbb{R}^p, \mu \geq 0} g(\mu,\lambda)}_{\text{best lower bound that } g \text{ gives on } p^*} \qquad (DP)$$

· The original problem (*) is called primal problem.

· (μ,λ) is called dual-feasible if $\mu \geq 0$ (pointwise) & $g(\mu,\lambda) > -\infty$

· (μ^*,λ^*) is called dual-optimal if it solves (DP).

Thm: (DP) is a convex optimisation problem. (no matter how (*) looks)

proof: $(DP) \Leftrightarrow \min_{(\tilde{\mu},\lambda) \in \mathbb{R}^m \times \mathbb{R}^p} -g(-\tilde{\mu},\lambda)$ s.t. $\tilde{\mu} \leq 0$

and $-g$ is convex. □

Example: dual problem to LP

The dual problem to a linear program is again an LP:

$$\min_x c^T x \quad \text{s.t.} \quad Ax - b = 0, \quad Bx - v \leq 0$$

$$g(\mu,\lambda) = \inf_x \; c x^T + \mu^T(Bx - v) + \lambda^T(Ax - b) = \begin{cases} -\mu^T v - \lambda^T b & \text{if } A^T\lambda + B^T\mu + c = 0 \\ -\infty & \text{else} \end{cases}$$

dual problem: $\max_{\mu \geq 0, \lambda} g(\mu,\lambda) = \max_{\substack{\mu \geq 0 \\ A^T\lambda + B^T\mu + c = 0}} -\mu^T v - \lambda^T b$

$$\Leftrightarrow \quad \min_{(\tilde{\mu},\lambda)} -v^T\tilde{\mu} + b^T\lambda \quad \text{s.t.} \quad \tilde{\mu} \leq 0, \; A^T\lambda - B^T\tilde{\mu} + c = 0$$

Strong versus weak duality

Let p^* be the optimum value of the primal pb. $(*)$, d^* of the dual pb. (DP).

We know $d^* \leq p^*$, in particular,

- if $(*)$ is unbounded from below ($p^* = -\infty$), then (DP) is infeasible ($d^* = -\infty$)
- if (DP) is unbounded from above ($d^* = \infty$), then $(*)$ is infeasible ($p^* = \infty$)

<u>Def</u>:
- The property $d^* \leq p^*$ is called <u>weak duality</u>.
- $p^* - d^* \geq 0$ is called <u>duality gap</u>.
- If $p^* = d^*$ we say that <u>strong duality</u> holds.

Typically, one does not have strong duality, however,
for many convex optimisation problems (and some non-convex ones) one does!
We will later consider criteria that imply strong duality.

In case of strong duality, $(*)$ and (DP) are equivalent!

Examples

Ex: $\min_{x \in \mathbb{R}} -x^2$ s.t. $x-1 \le 0$, $-x-1 \le 0$ (fo concave, unbdd below)

$L(x,\mu_1,\mu_2) = -x^2 + \mu_1(x-1) + \mu_2(-x-1)$

$g(\mu_1,\mu_2) = \inf_x L(x,\mu_1,\mu_2) = -\infty$

$\Rightarrow \quad d^* = -\infty < -1 = p^*$ (no strong duality)

Ex: $\min_{x,y \in \mathbb{R}} e^{-x}$ s.t. $\frac{x^2}{y} \le 0$, $5-y \le 0$ (convex problem, degenerate constraint)

$L(x,y,\mu_1,\mu_2) = e^{-x} + \mu_1 \frac{x^2}{y} + \mu_2(5-y)$

$g(\mu_1,\mu_2) = \inf_{x,y} L(x,y,\mu_1,\mu_2) = \begin{cases} 0 & \text{if } \mu_1=\mu_2=0 \\ -\infty & \text{else} \end{cases}$

$\Rightarrow \quad d^* = 0 < 1 = p^*$ (no strong duality)

Ex: $\min_{x,y \in \mathbb{R}} e^{-x}$ s.t. $x^2 \le 0$, $5-y \le 0$ (convex problem)

$L(x,y,\mu_1,\mu_2) = e^{-x} + \mu_1 x^2 + \mu_2(5-y)$

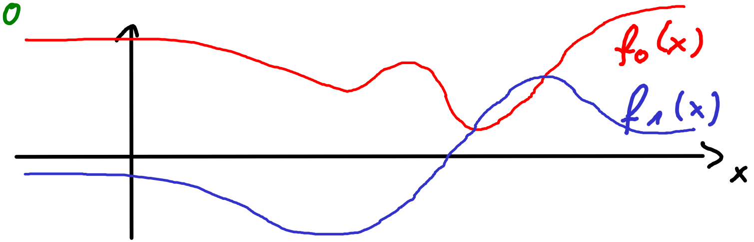
$g(\mu_1,\mu_2) = \begin{cases} -\infty & \text{if } \mu_2 \ne 0 \text{ or } \mu_1 < 0 \\ e^{-\hat{x}} + e^{-2\hat{x}}/4\mu_1 & \text{else, where } e^{-\hat{x}} = 2\mu_1 \hat{x} \end{cases}$

$\Rightarrow d^* = \sup_{\mu_1 \ge 0, \mu_2 \ge 0} g(\mu_1,\mu_2) = 1 = p^*$ (strong duality)

let $\mu_2 = 0$, $\mu_1 \to \infty$

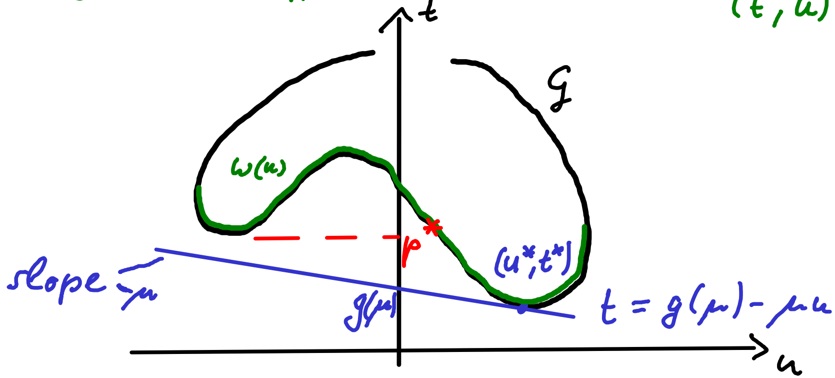
Geometric intuition and nonconvex problems

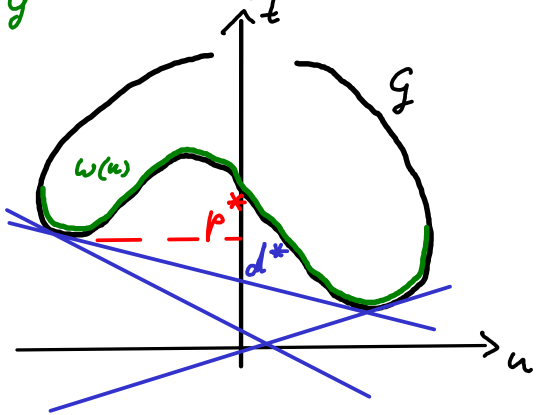
Ex: $\min_{x \in \mathbb{R}} f_0(x)$ s.t. $f_1(x) \le 0$



primal fcn: $\omega(u) = \inf \{ f_0(x) \mid f_1(x) = u \}$

dual fcn: $g(\mu) = \inf_x f_0(x) + \mu f_1(x) = \inf_{(t,u) \in \mathcal{G}} t + \mu u$ with $\mathcal{G} = \{(f_0(x), f_1(x)) \mid x \in \mathbb{R}\}$





if $g(\mu) = t^* + \mu u^*$, then $g(\mu)$ is the y-intercept of the line with slope $-\mu$ through (u^*, t^*)

strong duality only, if $\mathcal{G}$ lies on one side of tangent in x^*

Strong and weak duality of value sets

- As in the example, set $\mathcal{G} = \{(f_0(x), f_1(x), \dots, f_m(x), h_1(x), \dots, h_p(x)) \mid x \in \mathcal{D}\}$

  $\Rightarrow \; p^* = \inf \{ t \mid (t, u, v) \in \mathcal{G}, \; u \leq 0, \; v = 0 \}$

- Consider $(1, \mu, \lambda)^T (t, u, v) = t + \mu^T u + \lambda^T v$

  $\Rightarrow \; g(\mu, \lambda) = \inf \{ (1, \mu, \lambda)^T (t, u, v) \mid (t, u, v) \in \mathcal{G} \}$

  If infimum is finite, $(t, u, v) \longmapsto (1, \mu, \lambda)^T (t, u, v) = g(\mu, \lambda)$ defines a hyperplane orthogonal to $(1, \mu, \lambda)$ which is tangent to $\mathcal{G}$ and such that $\mathcal{G}$ lies on one side of it.

- $d^* = \sup\limits_{\lambda, \mu \geq 0} g(\mu, \lambda) = \sup\limits_{\lambda, \mu \geq 0} \inf \{ (1, \mu, \lambda)^T (t, u, v) \mid (t, u, v) \in \mathcal{G} \}$

  $\leq \inf \{ t \mid (t, u, v) \in \mathcal{G}, \; u \leq 0, \; v = 0 \} = p^*$

  if $\mu \geq 0$ & $u \leq 0, v = 0$, then $t \geq (1, \mu, \lambda)^T (t, u, v)$

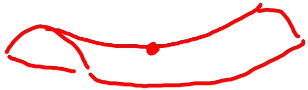
  $\Rightarrow$ weak duality, and strong duality iff $\mu^T u = 0$ at optimum

Saddle point interpretation

Def: Let $f : W \times Z \to \mathbb{R}$. $(\bar{w}, \bar{z}) \in W \times Z$ is called a <u>saddle point</u> if

$$f(\bar{w}, z) \le f(\bar{w}, \bar{z}) \le f(w, \bar{z}) \quad \forall (w, z) \in W \times Z, \text{ i.e.}$$

$$f(\bar{w}, \bar{z}) = \inf_{w \in W} f(w, \bar{z}) = \sup_{z \in Z} f(\bar{w}, z).$$



Thm: $$\sup_{z \in Z} \inf_{w \in W} f(w, z) \le \inf_{w \in W} \sup_{z \in Z} f(w, z)$$

proof: $$\sup_{z \in Z} \inf_{w \in W} f(w, z) \le \sup_{z \in Z} \inf_{w \in W} \sup_{\tilde{z} \in Z} f(w, \tilde{z}) = \inf_{w \in W} \sup_{\tilde{z} \in Z} f(w, \tilde{z}) \quad \square$$

Def: f possesses the <u>saddle point property</u> if $\sup_z \inf_w f(w,z) = \inf_w \sup_z f(w,z)$.

Thm: Strong duality $\Longleftrightarrow$ L satisfies the saddle point property on $\mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^p$

proof: $$p^* = \inf_{x,\, c_I(x) \le 0,\, c_E(x) = 0} f_0(x) = \inf_x \sup_{\mu \ge 0, \lambda} L(x, \mu, \lambda)$$

$$d^* = \sup_{\mu \ge 0, \lambda} g(\mu, \lambda) = \sup_{\mu \ge 0, \lambda} \inf_x L(x, \mu, \lambda) \qquad \square$$

$\Longrightarrow$ optimum is a saddle of L!

Certificates

<u>Notice</u>: if x feasible and (μ, λ) dual-feasible, then

$$f_0(x) - p^* \leq f_0(x) - g(\mu, \lambda) =: \varepsilon, \text{ i.e.}$$

x is ε-suboptimal.

In particular, if $f_0(x) = g(\mu, \lambda)$, then x is optimal.

In general, $p^*, d^* \in [g(\mu, \lambda), f_0(x)]$

<u>Def</u>: (μ, λ) is called a <u>certificate</u> that x is ε-suboptimal/optimal.

This can be exploited as stopping criterion for optimisation algorithms!

Complementarity

Assume strong duality, and let x^* primal-optimal and (μ^*, λ^*) dual-optimal

$$f_0(x^*) = p^* = d^* = g(\mu^*, \lambda^*) = \inf_x \{ f_0(x) + \mu^{*T} c_I(x) + \lambda^{*T} c_E(x) \}$$

$$\overset{(A)}{\leq} f_0(x^*) + \mu^{*T} c_I(x^*) + \lambda^{*T} c_E(x^*) \overset{(B)}{\underset{\mu^* \geq 0}{\leq}} f_0(x^*),$$

thus, equality holds in (A) and (B)

Equality in (A) $\Longrightarrow$ $x \mapsto L(x, \lambda^*, \mu^*)$ is minimised by x^* (saddle point property)

Equality in (B) $\Longrightarrow$ $\mu^{*T} c_I(x^*) = 0$ $\underset{\mu_i^* f_i(x^*) \leq 0}{\Longrightarrow}$ $\mu_i^* f_i(x^*) = 0$, $i = 1, \ldots, m$.

$\Longrightarrow$ $\mu_i^* = 0$ or $f_i(x^*) = 0$ (complementary slackness)

KKT conditions from strong duality

Thm: Consider $\min_{x \in \mathbb{R}^n} f_0(x)$ s.t. $c_I(x) \le 0$, $c_E(x) = 0$ with

f_0, c_E, c_I differentiable (i.e. $\mathcal{D} = \bigcap_{i=0}^{m} \operatorname{dom} f_i \cap \bigcap_{j=1}^{p} \operatorname{dom} h_j$ open).

If ~~strong duality holds~~ and x^* is primal-, (μ^*, λ^*) dual-optimal, then

before: regularity

$$\left.\begin{array}{l} c_I(x^*) \le 0,\ c_E(x^*) = 0,\ \mu^* \ge 0,\ \mu^{*T} c_I(x^*) = 0 \\ D f_0(x^*) + \mu^{*T} D c_I(x^*) + \lambda^{*T} D c_E(x^*) = 0 \end{array}\right\} \text{(KKT)}$$

proof: • saddle point property of L $\Rightarrow$ x^* minimises $x \mapsto L(x, \mu^*, \lambda^*)$

$$\Rightarrow 0 = \frac{\partial}{\partial x} L(x, \mu^*, \lambda^*) = D f_0(x^*) + \mu^{*T} D c_I(x^*) + \lambda^{*T} D c_E(x^*)$$

• complementary slackness $\Rightarrow$ $\mu^{*T} c_I(x^*) = 0$ $\square$

Thm: Let $f_0, \dots, f_m$ convex, $h_1, \dots, h_p$ affine.

Then (KKT) $\Rightarrow$ x^* is primal-, (μ^*, λ^*) dual-optimal.

proof: (KKT) implies feasibility of $x^*, (\mu^*, \lambda^*)$. Furthermore, due to $\mu^* \ge 0$,

$x \mapsto \underbrace{f_0(x) + \mu^{*T} c_I(x) + \lambda^{*T} c_E(x)}_{L(x, \mu^*, \lambda^*)}$ is convex $\overset{\text{(KKT)}}{\Longrightarrow}$ x^* minimises $L(x, \mu^*, \lambda^*)$

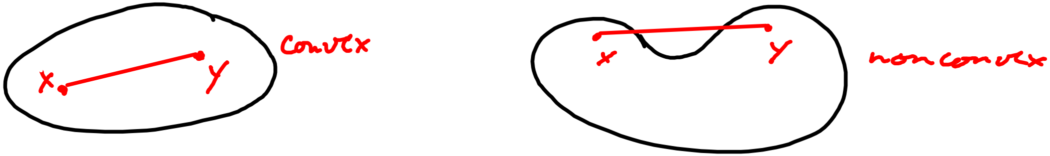
$$\Rightarrow g(\mu^*, \lambda^*) = L(x^*, \mu^*, \lambda^*) = f_0(x^*). \quad \square$$

Motivation

- Several practical optimisation problems are convex
  - → image processing / inverse problems (L^1-type optimisations)
  - → production planning: linearised constraints + bounds ⇒ LP
- Sometimes the convex relaxation (convex envelope) of a nonconvex optimisation problem can be found
- Convex optimisation problems are simpler to solve (at each point clear where to go next)
- Convex optimisation problems satisfy many fine properties
  - → local optimum = global optimum
  - → often strong duality
- Convex optimisation problems occur naturally, e.g. as dual problem.

Definition & examples of convex sets

__Def:__ A set $C \subset \mathbb{R}^n$ is called __convex__ if for all $x, y \in C$ the line segment in between is also contained in C, i.e. $\theta x + (1-\theta) y \in C \quad \forall x, y \in C, \theta \in [0,1]$



__Ex:__ • __Line:__ given $x, y \in \mathbb{R}^n$, $L = \{z \in \mathbb{R}^n \mid \exists\, \theta \in \mathbb{R} : z = \theta x + (1-\theta) y\}$ is convex

• __Affine sets:__ C is called affine if for all $x, y \in C$ also the line through x and y lies in C. If C is affine and $x_0 \in C$, then $V = \{z \mid z + x_0 \in C\}$ is a linear subspace.

→ e.g. the solution set to a linear equation, $C = \{x \mid Ax = b\}$, $A \in \mathbb{R}^{k \times n}$, $b \in \mathbb{R}^k$

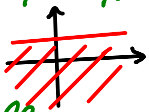
→ __affine hull__ of $C \subset \mathbb{R}^n$: $\operatorname{aff}(C) = \{y = \overset{\in \mathbb{R}}{\theta_1} x_1 + \ldots + \overset{\in \mathbb{R}}{\theta_k} x_k \mid k \in \mathbb{N}, \theta_1 + \ldots + \theta_k = 1, x_1, \ldots, x_k \in C\}$ is the smallest affine set containing C (homework)

→ __affine dimension__ of C = $\dim(\operatorname{aff}(C))$ (note: for S^1 this is 2!)

Examples II

<u>Ex</u>: · <u>Hyperplane</u>: $H = \{x \mid a^T x = b\}$ for $a \in \mathbb{R}^n, b \in \mathbb{R}$, is affine with $\operatorname{codim} H = 1$

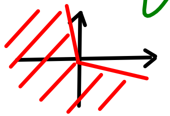
· <u>Halfspace</u>: $H = \{x \mid a^T x \leq b\}$, $a \in \mathbb{R}^n, b \in \mathbb{R}$, is convex

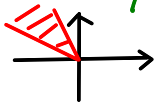


· <u>Cone</u>: C is called a cone if for all $x \in C$ and all $\theta \geq 0$ also $\theta x \in C$.

C is called a <u>convex cone</u> if it is additionally convex, i.e. $\forall x, y \in C$:

$\forall \theta_1, \theta_2 \geq 0 : \theta_1 x + \theta_2 y \in C$





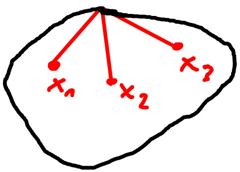
The <u>positive orthant</u> $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_1, \dots, x_n \geq 0\}$ is a convex cone.

· <u>(Norm-) ball</u>: Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$ (see a little later).

The (norm-)ball $\{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$ is convex

· <u>m-ellipse</u>: Let $x_1, \dots, x_m \in \mathbb{R}^n$. The m-ellipse

$\{x \mid |x - x_1| + \dots + |x - x_m| \leq 1\}$ is convex.



Convex combinations

Def: Let $x_1, \dots, x_k \in \mathbb{R}^n$, $\theta_1, \dots, \theta_k \geq 0$ with $\theta_1 + \dots + \theta_k = 1$. $y = \sum_{i=1}^{k} \theta_i x_i$ is called a convex combination of the x_i.

Thm: $C \subset \mathbb{R}^n$ is convex $\iff$ all convex combinations of any $x_1, \dots, x_k \in C$ are contained in C.

proof: "$\Leftarrow$" The line segment between $x, y \in C$ are exactly all convex comb. of x, y.

"$\Rightarrow$" induction over k: $k = 2$ is just definition of convexity

induction step: Consider $x_1, \dots, x_{k+1} \in C$, $\theta_1, \dots, \theta_{k+1} \geq 0$, $\theta_1 + \dots + \theta_{k+1} = 1$.

Wlog let $z = \sum_{i=1}^{k} \theta_i > 0$ and set $\theta_i' = \theta_i / z$, then $\sum_{i=1}^{k} \theta_i' = 1$.

Thus, $y' = \sum_{i=1}^{k} \theta_i' x_i \in C$.

Furthermore, $z + \theta_{k+1} = 1$, i.e. $\sum_{i=1}^{k+1} \theta_i x_i = z y' + \theta_{k+1} x_{k+1} \in C$,

since it can be written as convex combination of 2 elements. $\square$

Short: Convex combinations of convex combinations are again convex comb.

Convex hull

Thm: Let $\{C_i\}_{i\in J}$ be a family of convex sets, then $\bigcap_{i\in J} C_i$ is convex.

proof: Let $x,y \in \bigcap_{i\in J} C_i$, then $\forall i \in J$ we have $x,y \in C_i$, thus $\theta x + (1-\theta) y \in C_i \ \forall \theta \in [0,1]$

$\Rightarrow \forall \theta \in [0,1], \quad \theta x + (1-\theta) y \in \bigcap_{i\in J} C_i.$ $\square$

Def: The convex hull $\operatorname{conv} C$ of a set $C \subset \mathbb{R}^n$ is the intersection of all convex sets containing C.

Thm: $\operatorname{conv} C = \{ y = \sum_{i=1}^{k} \theta_i x_i \mid k \in \mathbb{N}, x_1, \dots, x_k \in C, \theta_1, \dots, \theta_k \geq 0, \sum_{i=1}^{k} \theta_i = 1 \} =: T$

proof: "$\subset$": T is convex: Any $x,y \in T$ are convex combs. of points in C, thus any conv. comb. of x and y is so, too, and thereby lies in T.

Also, $C \subset T$, thus $\operatorname{conv} C \subset T$.

"$\supset$": Let S be convex with $S \supset C$.

By the previous slide, S contains all its convex combinations, in particular also the convex combinations of C.

$\Rightarrow S \supset T \xRightarrow[\text{previous thm}]{} \operatorname{conv} C \supset T.$ $\square$

Caratheodory's theorem

Thm (Caratheodory): Every $x \in \operatorname{conv} G$, $G \subset \mathbb{R}^n$, can be written as convex combination of $n+1$ elements from G.

proof: Consider an arbitrary convex combination $x = \sum_{i=1}^{k} \theta_i x_i$ for $k > n+1$.

Claim: Without changing x one can change the θ_i s.t. one θ_i is 0.

Indeed, $\{x_2 - x_1, \dots, x_k - x_1\} \subset \mathbb{R}^n$ are linearly dependent, since $k-1 > n$.

$\Longrightarrow$ There are $(\beta_2, \dots, \beta_k) \in \mathbb{R}^{k-1} \setminus \{0\}$ s.t. $0 = \sum_{i=2}^{k} \beta_i (x_i - x_1) = \sum_{i=2}^{k} \beta_i x_i - \underbrace{\left(\sum_{i=2}^{k} \beta_i\right)}_{=:-\beta_1} x_1$.

Define $\theta_i' = \theta_i - t^* \beta_i$ for $t^* = \theta_{i^*}/\beta_{i^*}$, $i^* = \operatorname{argmin}_{i=1,\dots,k} \theta_i / |\beta_i|$

$\Longrightarrow$ $\theta_i' \geq 0$ and $\theta_i' = 0$ for at least one i.

Furthermore, $\sum_{i=1}^{k} \theta_i' = \underbrace{\sum_{i=1}^{k} \theta_i}_{1} - t^* \underbrace{\sum_{i=1}^{k} \beta_i}_{0} = 1$

and $\sum_{i=1}^{k} \theta_i' x_i = \underbrace{\sum_{i=1}^{k} \theta_i x_i}_{=x} - t^* \underbrace{\sum_{i=1}^{k} \beta_i x_i}_{= \sum_{i=2}^{k} \beta_i (x_i - x_1) = 0} = x$. $\square$

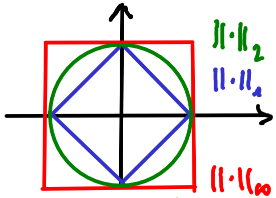
$\rightsquigarrow$ similar technique in simplex method!

Examples III: norm balls

Def: A map $\|\cdot\|: \mathbb{R}^n \to [0,\infty)$ is called a norm if for all $x, y \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$

(a) $\|x\| \geq 0$ (b) $\|x\| = 0 \Leftrightarrow x = 0$ (c) $\|\alpha x\| = |\alpha| \, \|x\|$ (d) $\|x+y\| \leq \|x\| + \|y\|$

Ex:
- Euclidean norm $\|x\|_2 = |x| = \sqrt{x^T x} = \sqrt{\sum_{i=1}^{n} x_i^2}$
- ℓ_1-norm $\|x\|_1 = \sum_{i=1}^{n} |x_i|$ "Manhattan norm"
- ℓ_∞-norm $\|x\|_\infty = \max_{i=1,\dots,n} |x_i|$



Def:
- The norm-ball $B(r,x) = \{y \in \mathbb{R}^n \mid \|x-y\| \leq r\}$ is convex. (homework)
- The ellipsoid $B_P(r,x) = \{y \in \mathbb{R}^n \mid (x-y)^T P (x-y) \leq r^2\}$, P pos. def. is convex.

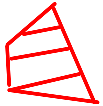
  $\|x\|_P = \sqrt{x^T P x}$ defines a norm.
- The norm cone $C = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} \mid \|x\| \leq t\}$ is a convex cone. "2^{nd} order cone" for $\|\cdot\|_2$

Examples IV: simplices

<u>Def</u>: • The <u>polyhedron</u> $P = \{x \in \mathbb{R}^n \mid a_j^T x \leq b_j,\ j = 1, \dots, m,\ c_j^T x = d_j,\ j = 1, \dots, p\}$, $a_j, c_j \in \mathbb{R}^n$, $b_j, d_j \in \mathbb{R}$

is the intersection of m half spaces and p hyperplanes and thus is convex.

Compact notation: $P = \{x \mid Ax \leq b,\ Cx = d\}$



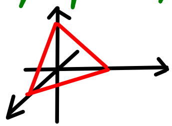
• A <u>polytope</u> is a bounded polyhedron.

• $k+1$ points $x_0, \dots, x_k \in \mathbb{R}^n$ are called <u>affinely independent</u>

if $x_1 - x_0, \dots, x_k - x_0$ are linearly independent.

• A <u>simplex</u> C is the convex hull of $k+1$ affinely indep. points, $C = \operatorname{conv}\{x_0, \dots, x_k\}$.

• The <u>probability simplex</u> is $\operatorname{conv}\{e_1, \dots, e_n\} \subset \mathbb{R}^n$



<u>Thm</u>: Every simplex is a polyhedron.

Topological notions

Def: • The closed convex hull $\overline{\text{conv}}\, S$ of a set $S \subset \mathbb{R}^n$ is the intersection of all closed convex sets containing S

• Let $C \subset \mathbb{R}^n$ be convex and non-empty. If the interior of C is non-empty, aff $C = \mathbb{R}^n$. The relative interior of $C \subset \mathbb{R}^n$, relint C, is the interior of C wrt the topology of aff C

$$x \in \text{relint } C \iff x \in \text{aff } C \text{ and } \exists \delta > 0: \text{aff } C \cap B(x,\delta) \subset C$$

• The relative boundary relbdy C is defined analogously.

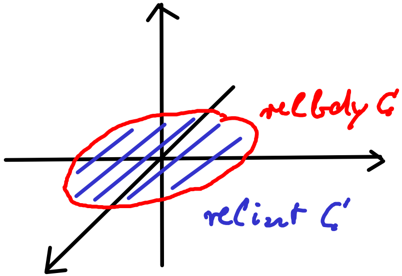
Ex: •

$$C = \{x \in \mathbb{R}^3 \mid x_3 = 0,\ x_1^2 + x_2^2 \leq 1\}$$

$$\text{aff } C = \{x \in \mathbb{R}^3 \mid x_3 = 0\}$$

$$\text{relint } C = \{x \in \mathbb{R}^3 \mid x_1^2 + x_2^2 < 1,\ x_3 = 0\}$$

$$\text{relbdy } C = \{x \in \mathbb{R}^3 \mid x_1^2 + x_2^2 = 1,\ x_3 = 0\}$$



Operations preserving convexity

Thm: The intersection of convex sets is convex (see earlier)

Thm: The cartesian product $C = C_1 \times \ldots \times C_k$ of convex sets $C_i \subset \mathbb{R}^{n_i}$ is convex.

proof: homework

Thm: Let $A : \mathbb{R}^n \to \mathbb{R}^m$ affine, i.e. $A(x) = Bx + c$ for $B \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^m$. Let $C \subset \mathbb{R}^n$, $D \subset \mathbb{R}^m$ convex

Then $A(C)$ and $A^{-1}(D) = \{ x \in \mathbb{R}^n \mid A(x) \in D \}$ are convex.

proof:
- Let $x, y \in C$. The image of $[x,y] = \{ \theta x + (1-\theta) y \mid \theta \in [0,1] \}$ is $[A(x), A(y)] \Rightarrow A(C)$ convex.
- Let $x, y \in A^{-1}(D)$, then $A([x,y]) = [A(x), A(y)] \subset D. \Rightarrow [x,y] \subset A^{-1}(D) \Rightarrow A^{-1}(D)$ convex. $\square$

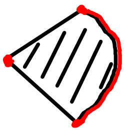
Thm: $C \subset \mathbb{R}^n$ convex $\Rightarrow$ $\operatorname{int} C$, $\operatorname{relint} C$, $\overline{C}$ convex.

proof: homework

Extreme points

Def: • Let $C \subset \mathbb{R}^n$ convex, $C \neq \emptyset$. $x \in C$ is called <u>extreme point</u> if there are no $x_1, x_2 \in C$, $x_1 \neq x_2$, with $x = \theta x_1 + (1-\theta) x_2$ for a $\theta \in (0,1)$.

• Set of extreme points $= \operatorname{ext} C$



Cor: • One may set $\theta = \frac{1}{2}$ in above definition.

• $x \in \operatorname{ext} C \iff C \setminus \{x\}$ convex

Ex: • $C = B(0,1) = \{x \in \mathbb{R}^n \mid |x| \leq 1\} \Longrightarrow$ all $x \in \mathbb{R}^n$ with $|x| = 1$ are extreme points:

Let $x = \theta x_1 + (1-\theta) x_2$, $\theta = \frac{1}{2}$, $x_1 \neq x_2$, $x_1, x_2 \in C$.

$$1 = |x|^2 = |\theta x_1 + (1-\theta) x_2|^2 = \underbrace{2\left(\theta^2 |x_1|^2 + (1-\theta)^2 |x_2|^2\right)}_{\leq 2\theta^2 + 2(1-\theta)^2 = 1} - \underbrace{|\theta x_1 - (1-\theta) x_2|^2}_{=0 \text{ iff } x_1 = \frac{1-\theta}{\theta} x_2 = x_2} < 1.$$

• The ℓ_1-ball has finitely many extreme points.

• $\operatorname{ext} C = \{0\}$ for a convex cone C which is not a halfhyperplane

• $\operatorname{ext} C = \emptyset$ for affine sets or halfspaces C.

Properties of extreme points

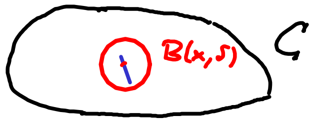
Thm: If $C \neq \emptyset$ compact and convex, then $\operatorname{ext} C \neq \emptyset$.

proof: Due to compactness, $x \mapsto |x|^2$ attains its maximum on C at some $\bar{x}$.

Claim: $\bar{x} \in \operatorname{ext} C$. Indeed, let $\bar{x} = \frac{x_1 + x_2}{2}$ for $x_1, x_2 \in C$, $x_1 \neq x_2$, then

$$|\bar{x}|^2 = \left|\frac{x_1+x_2}{2}\right|^2 = \frac{1}{2}\left(|x_1|^2 + |x_2|^2\right) - \left|\frac{x_1 - x_2}{2}\right|^2 < |\bar{x}|^2 \quad ↯. \qquad \square$$

Thm: $\operatorname{ext} C \subset \operatorname{relbdy} C$.



Thm (Minkowski): If $C \neq \emptyset$ compact and convex, then $C = \operatorname{conv}(\operatorname{ext} C)$.

proof: later...

Cor (see Caratheodory's Thm): If in addition $\operatorname{affdim} C = k$, then every point in C is the convex combination of at most $k+1$ extreme points.

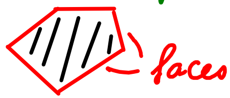
Faces

Def: A nonempty convex $F \subset C$ is called a <u>face</u> of C if for any $x_1, x_2 \in C$,

$$\operatorname{relint}[x_1, x_2] \cap F \neq \emptyset \quad \Longrightarrow \quad [x_1, x_2] \subset F.$$

Ex:

- $x \in \operatorname{ext}(C) \iff \{x\}$ is a (0-dimensional) face of C
- 1-dimensional faces = edges



- faces of polyhedra
- C itself is a face for C closed and convex

Thm: Let F be a face of the convex $C \subset \mathbb{R}^n$. $\operatorname{ext}(F) \subset \operatorname{ext}(C)$

proof: Assume $x \in \operatorname{ext} F \setminus \operatorname{ext} C$, i.e., there are $x_1, x_2 \in C$, $x_1 \neq x_2$, $x = \frac{x_1 + x_2}{2}$.

By definition of a face, $[x_1, x_2] \subset F$, i.e. $x_1, x_2 \in F \Longrightarrow x \notin \operatorname{ext} F$ ↯. $\square$

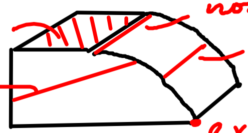
Exposed faces

Def: • A hyperplane $H_{s,r} = \{x \in \mathbb{R}^n \mid s^T x = r\}$ supports $C \subset \mathbb{R}^n$, if C completely lies in one of the halfspaces $\{x \in \mathbb{R}^n \mid s^T x \leq r\}$ or $\{x \in \mathbb{R}^n \mid s^T x \geq r\}$.

• $H_{s,r}$ supports C in $x \in C$, if in addition $x \in H_{s,r}$

• $F \subset C \subset \mathbb{R}^n$ is an exposed face if there is a supporting hyperplane H with $F = C \cap H$.

• Exposed point = 0-dimensional exposed face (corner)



Thm: An exposed face is a face.

proof: Let $F = C \cap H_{s,r}$ for a supporting hyperplane $H_{s,r}$ (wlog, $C \subset \{x \in \mathbb{R}^n \mid s^T x \leq r\}$).

Let $x, x_1, x_2 \in C$, $\alpha \in (0,1)$, with $x = \alpha x_1 + (1-\alpha) x_2 \in F \subset H_{s,r}$.

$\Rightarrow r = s^T(\alpha x_1 + (1-\alpha) x_2)$. Wlog assume $x_1 \notin F \subset H_{s,r}$, i.e. $s^T x_1 < r$.

$\Rightarrow s^T(\alpha x_1 + (1-\alpha) x_2) < r$ ↯ $\Rightarrow x_1, x_2 \in F \Rightarrow [x_1, x_2] \in F$. $\square$

Cor: For $C \subset \mathbb{R}^n$ convex, F an exposed face, $\mathrm{ext}(F) \subset \mathrm{ext}(C)$ (by previous slide).

Projections

Def: • A linear map $P : \mathbb{R}^n \to \mathbb{R}^n$ is called a <u>projection</u>, if $P^2 = P$ (idempotence).

• Let $V \subset \mathbb{R}^n$ be a subspace. $P_V : \mathbb{R}^n \to V$, $x \mapsto v$ for $x = v + v^\perp \in V \oplus V^\perp$ is called the <u>orthogonal projection</u> onto V.

Thm: P_V is linear, symmetric, positive semi-definite, idempotent, non-expansive ($\|P_V\|_2 \leq 1$), $x = P_V(x) + P_{V^\perp}(x)$.

How to generalise the orthogonal projection onto closed sets?

Thm: $P_V(x) = \operatorname*{argmin}_{y \in V} \frac{1}{2}\|y - x\|_2^2$.

proof: Let $x = v + v^\perp \in V \oplus V^\perp$, then $\frac{1}{2}\|y - x\|_2^2 = \frac{1}{2}\|y - v\|_2^2 + \frac{1}{2}\|v^\perp\|_2^2 \Rightarrow y = v$ is minimiser $\square$

Thm: Let $C \subset \mathbb{R}^n$ be closed, $x \in \mathbb{R}^n$, then $y \mapsto f_x(y) = \frac{1}{2}\|y - x\|_2^2$ attains its minimum on C.

proof: Let $c \in C$ and $S = \{y \in \mathbb{R}^n \mid f_x(y) \leq f_x(c)\}$. f_x is continuous, $C \cap S$ is compact

$\Rightarrow$ by Weierstraß' thm, f_x has a minimum on $C \cap S$ and thus on C. $\square$

Orthogonal projections

Def: Let $G \subset \mathbb{R}^n$ closed and convex. The (nonlinear) map $P_G : \mathbb{R}^n \to G$,

$$P_G(x) = \operatorname*{argmin}_{y \in G} \underbrace{\tfrac{1}{2}\|x-y\|_2^2}_{f_x(y)},$$ is called the orthogonal projection onto G.

Thm: P_G is well-defined.

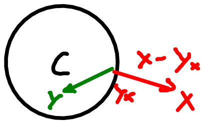
proof: Only uniqueness remains to be shown: Let $y_1 \neq y_2$ be minimisers, $y_0 = \frac{y_1+y_2}{2} \in G$.

$$f_x(y_0) - \tfrac{1}{2}\left(f_x(y_1) + f_x(y_2)\right) = \frac{\|x\|^2 - x\cdot y_1 - x\cdot y_2 + \|y_0\|^2}{2} - \frac{\|x\|^2 - 2x\cdot y_1 + \|y_1\|^2 + \|x\|^2 - 2x\cdot y_2 + \|y_2\|^2}{4}$$

$$= \tfrac{1}{8}\left(-\|y_1\|^2 - \|y_2\|^2 + 2y_1\cdot y_2\right) = -\tfrac{1}{8}\|y_1 - y_2\|^2 < 0 \ \lightning. \quad \square$$

Thm: • $P_G \circ P_G = P_G$ • P_G linear $\iff$ G is a subspace of $\mathbb{R}^n$.

Thm: $y_x = P_G(x) \iff (x - y_x)\cdot(y - y_x) \le 0 \quad \forall y \in G$.



proof: "$\Rightarrow$" For $y \in G$, $\alpha \in [0,1]$ arbitrary, consider $y_\alpha = y_x + \alpha(y - y_x) \in G$. We have

$$\tfrac{1}{2}\|y_x - x\|^2 = f_x(y_x) \le f_x(y_\alpha) = \tfrac{1}{2}\|y_x - x + \alpha(y - y_x)\|^2$$

$$\Rightarrow \quad 0 \le \alpha(y_x - x)\cdot(y - y_x) + \frac{\alpha^2}{2}\|y - y_x\|^2;$$ now divide by α and let $\alpha \to 0$.

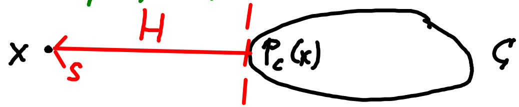
"$\Leftarrow$" $\forall y \in G$: $0 \ge (x - y_x)\cdot(y - y_x) = \|x - y_x\|^2 + (x - y_x)\cdot(y - x) \ge \|x - y_x\|^2 - \|x - y_x\|\,\|y - x\|$

$\Rightarrow$ either $y_x = x$ or $\|x - y_x\| \le \|y - x\|$. $\square$

Separation of convex sets

Thm: Let $C \subset \mathbb{R}^n$ convex, closed, $x \notin C$. There exists $s \in \mathbb{R}^n$ s.t. $\underbrace{s \cdot x > \sup_{y \in C} s \cdot y}_{\text{half space } H}$.

proof: Set $s := x - P_C(x)$.



For all $y \in C$, $0 \geq (x - P_C(x)) \cdot (y - P_C(x)) = s \cdot (y - x + s) = s \cdot y - s \cdot x + \|s\|^2 \Rightarrow s \cdot y < s \cdot x$ $\square$

Cor: Let $C_1, C_2 \subset \mathbb{R}^n$ closed, convex, nonempty, disjoint. If C_2 is bounded, there exists $s \in \mathbb{R}^n$ s.t. $\sup_{y \in C_1} s \cdot y < \min_{y \in C_2} s \cdot y$.

proof:
- $C_1 - C_2 = \{ y_1 - y_2 \mid y_1 \in C_1, y_2 \in C_2 \}$ is convex: indeed, $C_1 \times C_2$ is convex, and $C_1 - C_2 = i(C_1 \times C_2)$ for the linear $i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, $(a,b) \mapsto a-b$
- $C_1 - C_2$ is closed, since C_1 is closed and C_2 compact: indeed, let $y_k = y_k^1 - y_k^2 \in C_1 - C_2$ with $y_k \to y$. C_2 compact $\underset{\text{up to subsequence}}{\Rightarrow}$ $y_k^2 \to y^2 \in C_2$
$\Rightarrow y_k^1 = y_k + y_k^2 \to y + y^2 \in C_1$ (since C_1 closed) $\Rightarrow y \in C_1 - C_2$
- Previous thm for $x = 0$, $C = C_1 - C_2 \Rightarrow \exists s \in \mathbb{R}^n$: $0 = s \cdot 0 > \sup_{y \in C_1 - C_2} s \cdot y = \sup_{y_1 \in C_1} s \cdot y_1 - \min_{y_2 \in C_2} s \cdot y_2$ $\square$

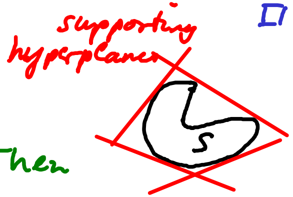
Note: · Statement can be wrong for C_1, C_2 unbounded

· if C_1, C_2 only convex, nonempty, disjoint, one still gets $\sup_{C_1} s \cdot y \leq \inf_{C_2} s \cdot y$ (homework)

Consequences of separation property: Supporting hyperplanes

Thm (existence of supporting hyperplane): Let $C \subset \mathbb{R}^n$ convex, $x \in \partial C \cap C$. There exists a supporting hyperplane to C in x,

proof: Consider a sequence $x_k \notin \overline{C}$ with $x_k \to x$. For each x_k there is $s_k \in \mathbb{R}^n$ with $s_k \cdot (x_k - y) > 0 \quad \forall y \in C$, Wlog, $\|s_k\|_2 = 1 \implies$ for a subsequence, $s_k \to s \in S^{n-1}$. By continuity $s \cdot (x - y) \geq 0 \quad \forall y \in C$. $\implies H_{s,r} = \{ y \mid s \cdot y = r \}$ is the sought hyperplane for $r = s \cdot x$. $\square$



Thm: For $S \subset \mathbb{R}^n$ and halfspaces $H^-_{s,r} = \{ y \in \mathbb{R}^n \mid s \cdot y \leq r \}$, set $\Sigma_S = \{ (s,r) \in \mathbb{R}^n \times \mathbb{R} \mid S \subset H^-_{s,r} \}$ and $C_S = \bigcap_{(s,r) \in \Sigma_S} H^-_{s,r}$. Then either $C_S = \overline{\mathrm{conv}}\, S$ or $\overline{\mathrm{conv}}\, S = \mathbb{R}^n$.

proof:
- Assume $\overline{\mathrm{conv}}\, S \neq \mathbb{R}^n$, then $C_S \supset \overline{\mathrm{conv}}\, S$.
- Now let $x \notin \overline{\mathrm{conv}}\, S$. Separate $\{x\}$ and $\overline{\mathrm{conv}}\, S$ by a hyperplane H_{s_0, r_0}, i.e.

$$s_0 \cdot x > \sup_{y \in S} s_0 \cdot y = -r_0 \,. \implies (s_0, r_0) \in \Sigma_S, \text{ but } x \notin H^-_{s_0,r_0} \implies x \notin C_S \,. \qquad \square$$

Consequences of separation property: halfspaces & extr. points

<u>Cor</u>: Let $C \subsetneq \mathbb{R}^n$ closed convex, then $C = C_C$, i.e. C is intersection of halfspaces.

In particular, a polyhedron is the intersection of finitely many halfspaces.

<u>Thm</u> (Minkowski): If $C \subset \mathbb{R}^n$ compact, convex, nonempty, then $C = \operatorname{conv}(\operatorname{ext} C)$.

proof: Induction in $k = \operatorname{affdim} C$; case $k = 0$ is trivial. Induction step $k-1 \to k$:

Let $x \in C$. • <u>case $x \in \operatorname{relbdy} C$</u>: There is a hyperplane $H \not\supset C$ supporting C in x. why?

$\Rightarrow$ $C \cap H$ has affine dimension $\leq k-1$

$\Rightarrow$ x is convex combination of extreme points x_i in $C \cap H$

x_i are also extreme points of C, since $C \cap H$ is an exposed face.

• <u>case $x \in \operatorname{relint} C$</u>: Choose $x' \neq x$, $x' \in C$

$\Rightarrow$ line through x, x' intersects $\operatorname{relbdy} C$ in two points y, z.

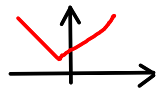
$\Rightarrow$ by first case, $y, z \in \operatorname{conv}(\operatorname{ext} C)$

$\Rightarrow$ $x \in \operatorname{conv}(\operatorname{ext} C)$ $\square$

Basic notions

Def: • Let $C \subset \mathbb{R}^n$ convex. $f : C \to \mathbb{R} \cup \{\infty\}$ is called convex ($f \in \mathrm{Conv}\, C$), if

$$\forall x, y \in C, \theta \in [0,1] : \quad f(\theta x + (1-\theta) y) \leq \theta f(x) + (1-\theta) f(y).$$



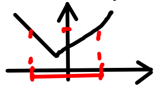
• f is strictly convex, if $\forall x \neq y \in C, \theta \in (0,1)$: $f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta) f(y)$

• f is concave if $-f$ is convex.

• The graph of f is the set $\{(x, f(x)) \mid x \in \mathrm{dom}\, f\}$.

• The epigraph of f is $\mathrm{epi}\, f = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq r\}$.

• The sublevel set of f wrt r is $S_r(f) = \{x \in \mathbb{R}^n \mid f(x) \leq r\}$.



Thm: • f convex $\Rightarrow$ $\mathrm{dom}\, f$ convex

• f convex $\Leftrightarrow$ $\mathrm{epi}\, f \subset \mathbb{R}^n \times \mathbb{R}$ convex

• $(x, r) \in \mathrm{epi}\, f \Leftrightarrow x \in S_r(f)$

Examples & properties

Ex: • characteristic or indicator function of a convex $C \subset \mathbb{R}^n$

$$I_C(x) = \begin{cases} 0 & \text{if } x \in C \\ \infty & \text{else} \end{cases} \qquad \text{different from } \chi_C = \begin{cases} 1 & \text{if } x \in C \\ 0 & \text{else} \end{cases} \quad !$$

• Any norm on $\mathbb{R}^n$ is convex: $\forall x, y \in \mathbb{R}^n, \theta \in (0,1): \|\theta x + (1-\theta) y\| \leq \|\theta x\| + \|(1-\theta) y\| = \theta \|x\| + (1-\theta)\|y\|$

• The maximum function $f(x) = \max_i x_i$ is convex.

• Linear and affine functions are convex; their epigraph is a halfspace.

Thm (Jensen's inequality): $\forall f \in \operatorname{Conv} \mathbb{R}^n, x_1, \dots, x_k \in \operatorname{dom} f, \theta_1, \dots, \theta_k \geq 0, \sum_{i=1}^k \theta_i = 1: f\left(\sum_{i=1}^k \theta_i x_i\right) \leq \sum_{i=1}^k \theta_i f(x_i)$

proof: $(x_i, f(x_i)) \in \operatorname{epi} f \xRightarrow[\text{epi } f \text{ convex}]{} \sum_{i=1}^k \theta_i (x_i, f(x_i)) = \left(\sum_{i=1}^k \theta_i x_i, \sum_{i=1}^k \theta_i f(x_i)\right) \in \operatorname{epi} f.$ □

Rem: Jensen's inequality also holds for integrals: Let $f \in \operatorname{Conv} \Omega$ and (Ω, Σ, P) be a probability space, then

$$f\Big(\underbrace{\int_\Omega x \, dP}_{\text{expectation of random variable } x}\Big) \leq \underbrace{\int_\Omega f(x) \, dP}_{\text{exp. of } f(x)}.$$

Convexity and continuity

necessary?!

Thm: Let $f \in \mathrm{Conv}\,\mathbb{R}^n$. f is continuous on $\mathrm{int}(\mathrm{dom} f)$.

proof: • Let $\bar{x} \in \mathrm{int}\,\mathrm{dom} f$; f is bounded in a neighbourhood of $\bar{x}$.

Indeed, there is a simplex $Q = \mathrm{conv}\{x_0, \dots, x_n\} \subset \mathrm{dom} f$ with $\bar{x} \in \mathrm{int}\, Q$.

Let $x \in Q$, i.e. $x = \sum_{i=1}^{n} \theta_i x_i$, $\theta_i \in [0,1]$, $\sum_{i=1}^{n} \theta_i = 1$, then $f(x) \le \sum_{i=1}^{n} \theta_i f(x_i) \le \max_i f(x_i) =: \tilde{M}$.

Also, for $\alpha = \frac{\mathrm{dist}(\bar{x}, \partial Q)}{2\,\mathrm{diam}(Q)} \in (0, \frac{1}{2})$ set $\hat{x} = \bar{x} - \frac{\alpha}{1-\alpha}(x - \bar{x}) \in Q$ (hence $\bar{x} = \alpha x + (1-\alpha)\hat{x}$)

$\Rightarrow f(\bar{x}) \le \alpha f(x) + (1-\alpha) f(\hat{x}) \Rightarrow f(x) \ge \frac{f(\bar{x}) - (1-\alpha) f(\hat{x})}{\alpha} \ge \min\left(-\tilde{M}, \frac{f(\bar{x}) - \tilde{M}}{\alpha}\right) =: -M$

• Let $v \in \mathbb{R}^n$ s.t. $\bar{x} + v \in Q$. For $\beta \in [0,1]$ we have

$$f(\bar{x} + \beta v) - f(\bar{x}) \le \beta f(\bar{x} + v) + (1-\beta) f(\bar{x}) - f(\bar{x}) \le \beta (M - f(\bar{x}))$$

• For $v \in \mathbb{R}^n$ small enough s.t. $\bar{x} \pm v \in Q$, $\beta \in [0,1]$ we have

$$f(\bar{x} + \beta v) - f(\bar{x}) \ge f(\bar{x} + \beta v) - \left[\frac{\beta}{1+\beta} f(\bar{x} - v) + \frac{1}{1+\beta} f(\bar{x} + \beta v)\right] = \frac{-\beta}{1+\beta}\left(f(\bar{x} + \beta v) - f(\bar{x} - v)\right) \ge \frac{-2M\beta}{1+\beta}$$

• Summarising, $|f(\bar{x} + \beta v) - f(\bar{x})| \le 2\beta M$ for all v small enough and $\beta \in [0,1]$

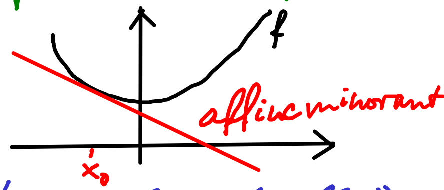
$\Rightarrow f$ is locally Lipschitz continuous. $\square$

Note: Not true in general ∞-dimensional spaces (e.g. $\|\cdot\|_{H^1} \in \mathrm{Conv}(H^1$ with norm $\|\cdot\|_{L^2})$; also there exist discontinuous linear functionals on ∞-dimensional Banach spaces)

Convex and affine functions

necessary ?!

<u>Thm</u>: Let $f \in \operatorname{Conv}\mathbb{R}^n$. $\forall x_0 \in \operatorname{relint} \operatorname{dom} f \;\; \exists s \in \mathbb{R}^n$: $f(x) \geq f(x_0) + s\cdot(x - x_0) \;\; \forall x \in \mathbb{R}^n$



proof:

- Note $\operatorname{aff} \operatorname{epi} f = (\operatorname{aff} \operatorname{dom} f) \times \mathbb{R}$ and $\operatorname{aff} \operatorname{dom} f = x_0 + V$ for a subspace $V \subset \mathbb{R}^n$.
- $(x_0, f(x_0)) \in \operatorname{relbdry}(\operatorname{epi} f) \implies \exists$ hyperplane supporting $\operatorname{epi} f$ in $(x_0, f(x_0))$, i.e. $\exists\, \tilde{s} \in V, \alpha \in \mathbb{R}, \begin{pmatrix}\tilde{s}\\ \alpha\end{pmatrix} \neq 0$ s.t. $\begin{pmatrix}\tilde{s}\\ \alpha\end{pmatrix}\cdot\begin{pmatrix}x\\ r\end{pmatrix} \leq \begin{pmatrix}\tilde{s}\\ \alpha\end{pmatrix}\cdot\begin{pmatrix}x_0\\ f(x_0)\end{pmatrix} \;\; \forall (x, r) \in \operatorname{epi} f$,
  in particular $-\alpha r \geq -\alpha f(x_0) + \tilde{s}\cdot(x - x_0)$ for all $x \in \operatorname{dom} f$, $r = f(x)$
- $\alpha < 0$. Indeed, letting $r \to \infty$ shows $\alpha \leq 0$,
  and $\alpha = 0$ implies $0 \geq \tilde{s}\cdot(x - x_0) \;\forall x \in \operatorname{dom} f$, i.e. $\tilde{s} = 0$, a contradiction.
- Dividing by $-\alpha$, $f(x) \geq f(x_0) + s\cdot(x - x_0)$ for $s = \dfrac{\tilde{s}}{-\alpha}$. $\square$

What is s if f is differentiable in x_0 ?

Differentiable convex functions

Thm: Let $C \subset \mathbb{R}^n$ open, convex, $f: C \to \mathbb{R}$ differentiable.

f is (strictly) convex on C $\iff$ $f(x) \geq (>) f(x_0) + \nabla f(x_0)\cdot(x - x_0)\ \forall x \neq x_0 \in C$

proof: "$\Rightarrow$": Let f be (strictly) convex, $\alpha \in (0,1)$, then

$$f(x_0 + \alpha(x - x_0)) - f(x_0) = f(\alpha x + (1-\alpha)x_0) - f(x_0) \overset{\leq}{<} \alpha f(x) + (1-\alpha)f(x_0) - f(x_0) = \alpha(f(x) - f(x_0))$$

Dividing by α and letting $\alpha \to 0$ yields $\nabla f(x_0)\cdot(x - x_0) \leq f(x) - f(x_0)$.

If equality holds and f is strictly convex, let $z = \theta x_0 + (1-\theta)x$, then

$$f(z) < \theta f(x_0) + (1-\theta) f(x) = f(x_0) + \nabla f(x_0)\cdot(x - x_0)(1-\theta) = f(x_0) + \nabla f(x_0)\cdot(z - x_0) \ \lightning$$

"$\Leftarrow$": Let $x_1 \neq x_2 \in C$, $\alpha \in (0,1)$, $x_0 = \alpha x_1 + (1-\alpha)x_2 \in C$.

$$(E_i) \qquad f(x_i) \overset{\geq}{>} f(x_0) + \nabla f(x_0)\cdot(x_i - x_0)$$

$$\alpha(E_1) + (1-\alpha)(E_2): \quad \alpha f(x_1) + (1-\alpha) f(x_2) \overset{\geq}{>} f(x_0) + \nabla f(x_0)\cdot(\underbrace{\alpha x_1 + (1-\alpha)x_2}_{x_0} - x_0)$$

$$\Rightarrow \quad f(\alpha x_1 + (1-\alpha)x_2) \overset{\leq}{<} \alpha f(x_1) + (1-\alpha) f(x_2) \qquad \square$$

Convexity and monotonicity of first derivative

Thm: Let $C \subset \mathbb{R}^n$ convex, $f: C \to \mathbb{R}$ differentiable. If Df is (strictly) monotone, i.e. $(Df(x) - Df(y)) \cdot (x-y) \geq (>)\ 0 \ \forall x \neq y \in C$, then f is (strictly) convex.

proof: Let $x_1, x_2 \in C$, $\alpha \in (0,1)$, $x = (1-\alpha)x_1 + \alpha x_2 = x_1 + \alpha(x_2 - x_1) \in C$, $x_s = x_1 + s(x_2 - x_1)$

$$f(x) = f(x_1) + \int_0^\alpha Df(x_s)(x_2 - x_1)\,ds \quad (E_1) \qquad f(x_2) = f(x) + \int_\alpha^1 Df(x_s)(x_2 - x_1)\,ds \quad (E_2)$$

$$(1-\alpha)(E_1) - \alpha(E_2) \Rightarrow \quad f(x) = (1-\alpha)f(x_1) + \alpha f(x_2) + \underbrace{\left[(1-\alpha)\int_0^\alpha Df(x_s)\,ds - \alpha\int_\alpha^1 Df(x_s)\,ds\right](x_2 - x_1)}_{A}$$

$$A = \alpha(1-\alpha)\int_0^1 Df(x_{\alpha s}) - Df(x_{\alpha + (1-\alpha)s})\,ds\,(x_2 - x_1) \leq (<)\ 0. \qquad \square$$

Thm: Let $C \subset \mathbb{R}^n$ convex, $f: C \to \mathbb{R}$ differentiable. f (strictly) convex $\Rightarrow$ Df (strictly) monotone.

proof: $f(x_1) \underset{(>)}{\geq} f(x_2) + Df(x_2)(x_1 - x_2)$ & $f(x_2) \underset{(>)}{\geq} f(x_1) + Df(x_1)(x_2 - x_1)$.

Now sum both inequalities. $\square$

Cor: Let $C \subset \mathbb{R}^n$ convex, $f: C \to \mathbb{R}$ twice differentiable. Then

(a) $D^2 f$ (strictly) positive definite $\Rightarrow$ f (strictly) convex $\quad$ (b) f convex $\Rightarrow D^2 f$ pos. semi-def

$\Leftarrow$ is false in general!

proof: (strict) positive definiteness of $D^2 f$ $\overset{\Longleftarrow}{\Longrightarrow}$ (strict) monotonicity of Df. $\square$

Operations preserving convexity

<u>Thm</u>: i) Let $f_1, \dots, f_m \in \operatorname{Conv} \mathbb{R}^n$, $t_1, \dots, t_m \geq 0$, $x_0 \in \mathbb{R}^n$ with $f_1(x_0), \dots, f_m(x_0) < \infty$.

Then $\sum_{i=1}^{m} t_i f_i \in \operatorname{Conv} \mathbb{R}^n$.

ii) Let $f \in \operatorname{Conv} \mathbb{R}^n$, $A : \mathbb{R}^n \to \mathbb{R}^n$ affine with $\operatorname{range} A \cap \operatorname{dom} f \neq \emptyset$.

Then $f \circ A \in \operatorname{Conv} \mathbb{R}^n$.

iii) Let $f_i \in \operatorname{Conv} \mathbb{R}^n$, $i \in \mathcal{J}$, $x_0 \in \mathbb{R}^n$ with $f_i(x_0) < M < \infty$ $\forall i \in \mathcal{J}$.

Then $x \mapsto \sup_{i \in \mathcal{J}} f_i(x) \in \operatorname{Conv} \mathbb{R}^n$.

iv) Let $f \in \operatorname{Conv} \mathbb{R}^n$, $h \in \operatorname{Conv} \mathbb{R}$ monotonically increasing, $x_0 \in \mathbb{R}^n$ with $f(x_0) \in \operatorname{dom} h$. $h \circ f \in \operatorname{Conv} \mathbb{R}^n$

v) Let $f \in \operatorname{Conv} \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, $C \subset \mathbb{R}^{n_2}$ convex, $g(x) = \inf_{y \in C} f(x,y)$. If $g(x) > -\infty$ $\forall x \in \mathbb{R}^n$, $g \in \operatorname{Conv} \mathbb{R}^{n_1}$

Example: Moreau-Yosida approximation for $f \in \operatorname{Conv} \mathbb{R}^n$, $\lambda > 0$:

$$f_\lambda(x) = \inf_{y \in \mathbb{R}^n} f(y) + \frac{1}{2\lambda} \|x - y\|_2^2 .$$

proof: i), ii), iv) homework

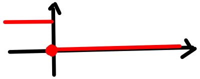
iii) Let $f(x) = \sup_{i \in \mathcal{J}} f_i(x)$. $\operatorname{epi} f = \bigcap_{i \in \mathcal{J}} \operatorname{epi} f_i$ is convex.

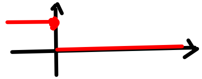
v) $\operatorname{epi} g = \{(x,t) \mid \exists y \in C : (x,y,t) \in \operatorname{epi} f\} = P_{\mathbb{R}^{n_1} \times \mathbb{R}}(\operatorname{epi} f)$ is convex. $\square$

Semi-continuity

Def: $f: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is called <u>lower semi-continuous</u> if $\liminf_{y \to x} f(y) \geq f(x) \; \forall x \in \mathbb{R}^n$.

Ex: $f(x) = \begin{cases} 1 & x < 0 \\ 0 & x \geq 0 \end{cases}$ is lower semi-continuous, $f(x) = \begin{cases} 1, & x \leq 0 \\ 0, & x > 0 \end{cases}$ is not.





Thm: Let $f: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$. The following are equivalent:

(a) f is lower semi-continuous

(b) $\operatorname{epi} f$ is closed in $\mathbb{R}^n \times \mathbb{R}$

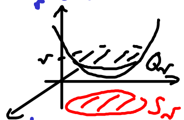
(c) The sublevel sets $S_r(f)$ are closed for all $r \in \mathbb{R}$.

proof: (a) $\Rightarrow$ (b): Let $(y_k, r_k) \in \operatorname{epi} f$, $(y, r) = \lim_{k \to \infty} (y_k, r_k)$.

$$r = \lim_{k \to \infty} r_k \geq \liminf_{k \to \infty} f(y_k) \geq \liminf_{x \to y} f(x) \geq f(y), \text{ hence } (y, r) \in \operatorname{epi} f.$$

(b) $\Rightarrow$ (c): Let $A_r(x) = (x, r)$, $Q_r = \operatorname{epi} f \cap (\mathbb{R}^n \times \{r\})$. Q_r is closed, A_r continuous.

$$S_r(f) = \{x \in \mathbb{R}^n \mid f(x) \leq r\} = A_r^{-1}(Q_r) \text{ is closed.}$$



(c) $\Rightarrow$ (a): Assume (a) is false, i.e. $\exists\, y_k \to x$ with $\rho := \lim_{k \to \infty} f(y_k) < f(x)$. Let $r \in (\rho, f(x))$.

For $k > k_0$ large enough, $f(y_k) \leq r < f(x)$, i.e. $y_k \in S_r(f) \; \forall k > k_0$, but $x \notin S_r(f)$ ↯ □

Closed convex functions

<u>Def</u>: • A function is called <u>closed</u> if it is lsc on $\mathbb{R}^n$ (or if epi f is closed).

• The <u>relaxation</u> (or <u>lower semi-continuous envelope</u>) of a function f is defined via $\bar{f}(x) := \mathrm{cl}\, f(x) := \liminf_{y \to x} f(y)$ or $\mathrm{epi}\, \bar{f} = \overline{\mathrm{epi}\, f}$

<u>Thm</u>: $f \in \mathrm{Conv}\, \mathbb{R}^n \Rightarrow \bar{f} \in \mathrm{Conv}\, \mathbb{R}^n$, and $f = \bar{f}$ on relint dom f.

proof: • f convex $\Rightarrow$ epi f convex $\Rightarrow$ $\mathrm{epi}\, \bar{f} = \overline{\mathrm{epi}\, f}$ convex $\Rightarrow \bar{f}$ convex.

Also, $\bar{f} \le f \not\equiv \infty$ and $\bar{f} > -\infty \ \forall x$ (due to minorisation by affine fcns.)

• $f|_{\mathrm{aff\,dom} f}$ is continuous on relint dom f,

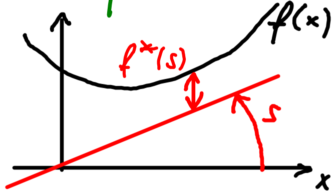
thus $\bar{f}|_{\mathrm{aff\,dom} f} = \overline{f|_{\mathrm{aff\,dom} f}} = f|_{\mathrm{aff\,dom}}$ on relint dom f

$\Rightarrow f = \bar{f}$ on relint dom f. $\square$

The conjugate function

<u>Def</u>.: The <u>conjugate function</u> (<u>Legendre-Fenchel dual</u>) to a function $f: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ (not necessarily convex) is defined as

$$f^*(s) = \sup_{x \in \mathbb{R}^n} s \cdot x - f(x), \quad s \in \mathbb{R}^n$$



- The map $f \mapsto f^*$ is the <u>Legendre-Fenchel-transform</u>

<u>Thm:</u> (Fenchel-Young-inequality) $f^*(y) + f(x) \geq x \cdot y \quad \forall x, y \in \mathbb{R}^n$

<u>Ex:</u> $f(x) = \begin{cases} x \log x - x, & x \geq 0 \\ \infty & \text{else} \end{cases}$, $f^*(y) = e^y$, Young-inequality $xy \leq x \log x - x + e^y \quad \forall x \geq 0, y \in \mathbb{R}$

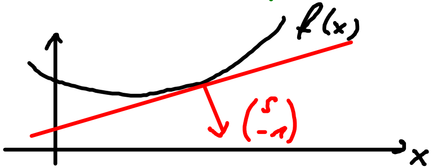
<u>Thm:</u> Let $f: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, $f \not\equiv \infty$, $f \geq g$ for an affine function g. f^* is closed convex.

proof:
- f^* is supremum of affine functions $y \mapsto f_x(y) = y \cdot x - f(x)$ $\Rightarrow f^*$ is convex
- $\operatorname{epi} f^* = \bigcap_{x \in \mathbb{R}^n} \operatorname{epi} f_x$ is closed as intersection of half spaces $\Rightarrow f^*$ is closed
- $f^* \not\equiv \infty$: Indeed, let $g(x) = s_0 \cdot x + r_0$, then $f^*(s_0) = \sup_x s_0 \cdot x - f(x) \leq \sup_x s_0 \cdot x - g(x) = -r_0$ $\square$

The biconjugate

Thm: Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, $f \not\equiv \infty$, $f(x) \geq s_0 \cdot x + r_0$ for some $(s_0, r_0) \in \mathbb{R}^n \times \mathbb{R}$.

We have $\operatorname{epi} f^{**} = \overline{\operatorname{conv}} \operatorname{epi} f$, i.e. f^{**} is largest closed convex function below f ("convex relaxation", "convex, lower semi-continuous envelope").



Cor: $f^{**} = f \iff f$ is closed convex

proof:
- Let $\Sigma = \{ (s,r) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \geq s \cdot x - r \;\; \forall x \}$ = parameters of supp. hyperplanes of epi f

  $\binom{s}{-1} \cdot \binom{x}{f(x)} \leq r$

- $f^{**}(x) = \sup\limits_{(s,r) \in \Sigma} (s \cdot x - r)$ = sup over all supp. hyperplanes of their vertical coordinate at x

  Indeed, $(s,r) \in \Sigma \iff r \geq \sup\limits_{x \in \mathbb{R}^n} s \cdot x - f(x) = f^*(s)$.

  Thus, $\sup\limits_{(s,r) \in \Sigma} (s \cdot x - r) = \sup\limits_{s,r \text{ with } f^*(s) \leq r} s \cdot x - r = \sup\limits_{s} s \cdot x - f^*(s) = f^{**}(x)$.

- $\operatorname{epi} f^{**} = \{ (x,t) \mid f^{**}(x) \leq t \} = \{ (x,t) \mid \sup\limits_{(s,r) \in \Sigma} (s \cdot x - r) \leq t \} = \bigcap\limits_{(s,r) \in \Sigma} \{ (x,t) \mid s \cdot x - r \leq t \}$

  = intersection of all half spaces containing epi f = $\overline{\operatorname{conv}} \operatorname{epi} f$ $\square$

Examples

Ex: · if f convex & differentiable, dom $f = \mathbb{R}^n$:

$$x^* = \operatorname{argmax}_x \; y \cdot x - f(x) \iff y = \nabla f(x^*) \qquad \Longrightarrow \quad f^*(y) = x^* \cdot \nabla f(x^*) - f(x^*)$$

· Let $Q \in \mathbb{R}^{n \times n}$ symmetric positive definite, $b \in \mathbb{R}^n$, $f(x) = \frac{1}{2} x^T Q x + b^T x$.

$$f^*(y) = \max_x \; y^T x - \frac{1}{2} x^T Q x - b^T x$$

$$y = \nabla f(x^*) = Qx^* + b \implies x^* = Q^{-1}(y - b) \implies f^*(y) = \frac{1}{2}(y-b)^T Q^{-1}(y-b)$$

· special case $Q = I$, $b = 0$, i.e. $f(x) = \frac{1}{2}\|x\|_2^2 \implies f^*(y) = \frac{1}{2}\|y\|_2^2$

· Fenchel's inequality: $x^T Q x + y^T Q^{-1} y \geq 2 x \cdot y$ (case $b = 0$)

$\|x\|^2 + \|y\|^2 \geq 2 x \cdot y$ (case $b = 0$, $Q = I$)

· $f(x) = I_C(x) = \begin{cases} 0 & \text{if } x \in C \subset \mathbb{R}^n \\ \infty & \text{else} \end{cases} \implies I_C^*(y) = \sup_{x \in C} y \cdot x =: \sigma_C(y)$

"indicator function" "support function"

· special case C = subspace of $\mathbb{R}^n$: $\sigma_C(y) = I_{C^\perp}(y)$

· $C = B(0,1) = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\} \implies \sigma_C(y) = \|y\|_*$ for the dual norm $\|y\|_* = \sup_{\|x\| \leq 1} |x \cdot y|$

· $\|\cdot\|_p$ & $\|\cdot\|_q$ with $\frac{1}{p} + \frac{1}{q} = 1$ are dual norms $\implies \sigma_{B(0,1), \|\cdot\|_p} = \|\cdot\|_q$

Calculation rules

Thm: Let $f, f_i : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, $f, f_i \not\equiv \infty$, $f, f_i \geq g, g_i$ for affine g, g_i

$h(x)$	$h^*(y)$
$f(x) + r$	$f^*(y) - r$
$t f(x)$, $t > 0$	$t f^*\left(\frac{y}{t}\right)$
$f(tx)$, $t \neq 0$	$f^*\left(\frac{y}{t}\right)$
$f(Ax)$, $A \in \mathbb{R}^{n \times n}$ invertible	$f^*(A^{-T} y)$
$f(x - x_0)$	$f^*(y) + y \cdot x_0$

$h(x)$	$h^*(y)$
$f(x) + y_0 \cdot x$	$f^*(y - y_0)$
$\sum_{j=1}^{m} f_j(x_j)$, $x = (x_1, \dots, x_m)$, $x_j \in \mathbb{R}^{n_j}$	$\sum_{j=1}^{m} f_j^*(y_j)$, $y = (y_1, \dots, y_m)$
$\leq f$	$\geq f^*$
$\alpha f_1 + (1-\alpha) f_2$, $\alpha \in [0,1]$	$\leq \alpha f_1^* + (1-\alpha) f_2^*$

$\underbrace{\qquad\qquad}_{\text{"convexity of conjugation"}}$

Def: · Given $f, g : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, their <u>infimal convolution</u> is

$$(f \,\square\, g)(x) = \inf_{y \in \mathbb{R}^n} f(y) + g(x-y) = \inf_{y \in \mathbb{R}^n} g(y) + f(x-y)$$

· <u>Moreau-Yosida-approximation</u> $\quad f_\lambda = f \,\square\, \frac{1}{2\lambda} \|\cdot\|_2^2$

Thm: $(f \,\square\, g)^* = f^* + g^*$ careful: if $f, g, f^* \square g^*$ closed convex, then $(f+g)^* = f^* \square g^*$, but not generally

proof: $(f \,\square\, g)^*(y) = \sup_x x \cdot y - \inf_{x_1 + x_2 = x} \big(f(x_1) + g(x_2)\big) = \sup_{x_1, x_2} (x_1 + x_2) \cdot y - f(x_1) - g(x_2) = f^*(y) + g^*(y)$ □

Coercivity and the conjugate

Def: · A function f is called coercive or 0-coercive if $\lim_{\|x\|\to\infty} f(x)=\infty$

· f is called 1-coercive if $\lim_{\|x\|\to\infty} \frac{f(x)}{\|x\|}=\infty$

Thm: Let $f:\mathbb{R}^n\to\mathbb{R}\cup\{\infty\}$, $f\not\equiv\infty$, $f(x)\geq y_0\cdot x+r_0$. f 1-coercive $\Longrightarrow f^*<\infty$

proof: Let $y\in\mathbb{R}^n$. (a) $\exists R>0: f(x)\geq\|y\|_2\|x\|_2 \ \forall \|x\|_2\geq R$

show $f^*(y)<\infty$!

$\underset{\text{Cauchy-Schwarz}}{\Longrightarrow}$ $x\cdot y-f(x)\leq 0 \ \forall \|x\|_2\geq R \Longrightarrow \sup_{\|x\|_2\geq R} x\cdot y-f(x)\leq 0$

(b) $\sup_{\|x\|_2<R} x\cdot y-f(x)\leq \sup_{\|x\|_2<R} x\cdot y-y_0\cdot x-r_0\leq R\|y-y_0\|_2-r_0$

(a) & (b) $\Longrightarrow f^*(y)\leq\max(0, R\|y-y_0\|-r_0)$ $\square$

Thm: $F_1:=\{\text{support functions of convex sets}\}=\{\text{closed convex positively homogeneous fcns}\}=:F_2$

proof: $\sigma\in F_1 \Longleftrightarrow \sigma=I_C^*$ for some convex $C\Longrightarrow\sigma$ closed convex with $\sigma(\lambda x)=|\lambda|\sigma(x) \ \forall\lambda\geq 0 \Longleftrightarrow \sigma\in F_2$

$\sigma\in F_2 \Longrightarrow \sigma^*(y)=\sup_x x\cdot y-\sigma(x)=\sup_{\lambda\geq 0,x}\lambda x\cdot y-\sigma(\lambda x)=\sup_{\lambda\geq 0}\lambda\,\sigma^*(y)=I_C(y)$

for $C=\{y\in\mathbb{R}^n \mid \sigma^*(y)=0\}$; C must be convex since σ^* is $\Rightarrow\sigma\in F_1$ $\square$

Cor: $f\in\operatorname{Conv}\mathbb{R}^n$ pos. homogeneous $\Rightarrow \operatorname{cl} f=\sigma_C$ for the closed convex $C=\{x\in\mathbb{R}^n \mid x\cdot y\leq f(y)\ \forall y\in\mathbb{R}^n\}$

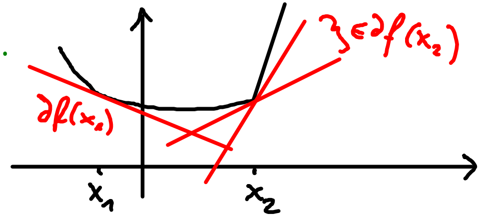
proof: · $\operatorname{cl} f\in F_2 \Rightarrow \operatorname{cl} f\in F_1 \Rightarrow \operatorname{cl} f=\sigma_C$ for some convex $C\subset\mathbb{R}^n$

· $f^*(x)=\sigma_C^*(x)=I_C(x)=\begin{cases}\infty & \text{if } \exists\, y\in\mathbb{R}^n: x\cdot y>f(y)\\ 0 & \text{else}\end{cases}$ $\square$

The subdifferential

Def: · The <u>subdifferential</u> of $f \in \operatorname{Conv}\mathbb{R}^n$ is $\partial f(x) = \{ s \in \mathbb{R}^n \mid f(y) \geq f(x) + s\cdot(y-x) \ \forall y \in \mathbb{R}^n \}$

· Elements of ∂f are called <u>subgradients</u>.



Cor: $\partial f(x) \neq \emptyset$ for all $x \in \operatorname{relint} \operatorname{dom} f$ (see earlier proof!)

Thm: $\partial f(x)$ is closed convex for all $x \in \operatorname{dom} f$ (for $x \notin \operatorname{dom} f$, $\partial f(x) = \emptyset$)

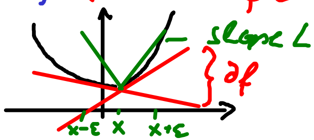
proof: $\partial f(x) = \bigcap_{y \in \mathbb{R}^n} \{ s \in \mathbb{R}^n \mid s\cdot(y-x) \leq f(y) - f(x) \}$ is intersection of halfspaces. □

Thm: $\partial f(x)$ is bounded for all $x \in$ int dom f (necessary?)

∂f is: · nonempty · compact · convex

proof: There is $\varepsilon > 0$ s.t. f is Lipschitz continuous on $B(x,\varepsilon)$ (recall: $f \in C^{0,1}_{loc}$!)

Let $y = x + \frac{\delta s}{\|s\|_2}$ for $s \in \partial f(x)$, $\delta \in (0,\varepsilon)$, then



$$f(x) + \underset{\uparrow}{L}\delta \geq f(y) \underset{\uparrow}{\geq} f(x) + s\cdot\frac{\delta s}{\|s\|_2} = f(x) + \delta\|s\|_2 \Rightarrow \|s\|_2 \leq L \ \square$$

Lipschitz constant $\qquad s \in \partial f(x)$

Subdifferential and directional derivatives

Def: Let $f \in \operatorname{Conv} \mathbb{R}^n$. The directional derivative in direction $v \in \mathbb{R}^n$ at $x \in \operatorname{dom} f$ is defined as

$$\partial_v f(x) = \liminf_{\lambda > 0,\, \lambda \to 0} \frac{f(x+\lambda v) - f(x)}{\lambda} = \inf_{\lambda > 0} \frac{f(x+\lambda v) - f(x)}{\lambda}$$

Rem: $\partial_{(\cdot)} f(x)$ for $f \in \operatorname{Conv} \mathbb{R}^n$ is convex by definition.

Thm: Let $f \in \operatorname{Conv} \mathbb{R}^n$, $f(x) < \infty$. $\quad s \in \partial f(x) \iff \partial_v f(x) \geq s \cdot v \quad \forall v \in \mathbb{R}^n$

· In fact, $\operatorname{cl}(\partial_{(\cdot)} f(x)) : \mathbb{R}^n \to \mathbb{R}$ is support function of closed convex $\partial f(x)$, $v \longmapsto \sup_{s \in \partial f(x)} s \cdot v$

proof: · $s \in \partial f(x) \iff \frac{f(x+\lambda v) - f(x)}{\lambda} \geq v \cdot s \;\; \forall \lambda > 0, v \in \mathbb{R}^n \iff \partial_v f(x) \geq s \cdot v \;\; \forall v \in \mathbb{R}^n$

· $\partial_{(\cdot)} f(x)$ is pos. homogeneous

$$\Rightarrow \operatorname{cl}(\partial_{(\cdot)} f(x)) = \sigma_C \text{ for } C = \{ s \in \mathbb{R}^n \mid s \cdot v \leq \partial_v f(x) \;\forall v \in \mathbb{R}^n \} = \partial f(x) \qquad \square$$

Thm: Let $f \in \operatorname{Conv} \mathbb{R}^n$. $\partial f(x)$ nonempty & bdd $\iff x \in \operatorname{int} \operatorname{dom} f$

proof: $\partial f(x) \neq \emptyset$ & bdd $\iff |\sigma_{\partial f(x)}| < \infty$ everywhere $\iff \operatorname{cl}(\partial_{(\cdot)} f(x)) < \infty$

$$\underset{\substack{\uparrow \\ \text{convexity of } \partial_{(\cdot)} f(x)}}{\iff} \partial_{(\cdot)} f(x) < \infty \iff x \in \operatorname{int} \operatorname{dom} f \qquad \square$$

Relation to differential

<u>Thm:</u> Let $f \in \operatorname{Conv}\mathbb{R}^n$. f differentiable in x with $\nabla f(x) = s \iff \partial f(x) = \{s\}$

proof: "$\Rightarrow$": Let $\tilde{s} \in \partial f(x)$, $v \in \mathbb{R}^n$.

$$
\begin{aligned}
\text{subdifferential}: \quad & f(x+v) \geq f(x) + \tilde{s} \cdot v \\
-\text{Taylor}: \quad & f(x+v) = f(x) + \nabla f(x) \cdot v + o(\|v\|) \\
\hline
& 0 \geq (\tilde{s} - \nabla f(x)) \cdot v + o(\|v\|) \quad \forall v \in \mathbb{R}^n
\end{aligned}
$$

"$\Leftarrow$": $(\operatorname{cl} \partial_{(\cdot)} f(x))(y) = \sigma_{\partial f(x)}(y) = s \cdot y$

$\Rightarrow \quad \partial_y f(x) = s \cdot y \quad \forall y$ (since for g convex $\operatorname{cl} g \neq g$ at most on rel bdy dom g)

$\Rightarrow \quad \nabla f(x) = s$ $\square$

<u>Thm:</u> Let $f \in \operatorname{Conv}\mathbb{R}^n$. $x^* = \operatorname{argmin}_x f(x) \iff 0 \in \partial f(x^*)$

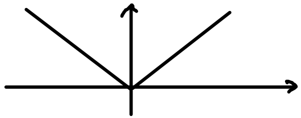
proof: $x^* = \operatorname{argmin}_x f(x)$

$\iff f(x) \geq f(x^*) \quad \forall x \in \mathbb{R}^n$

$\iff f(x) \geq f(x^*) + 0 \cdot (x - x^*) \quad \forall x \in \mathbb{R}^n$

$\iff 0 \in \partial f(x^*)$ $\square$

Examples and properties



Ex:

- $f(x) = |x|,\ x \in \mathbb{R} \Rightarrow \partial f(x) = \begin{cases} \{\operatorname{sign} x\}, & x \neq 0 \\ [-1,1], & x = 0 \end{cases}$
- $f(x) = \|x\|_2,\ x \in \mathbb{R}^n \Rightarrow \partial f(x) = \begin{cases} \{x/\|x\|_2\}, & x \neq 0 \\ B_{\|\cdot\|_2}(0,1), & x = 0 \end{cases}$
- $f(x) = \|x\|,\ x \in \mathbb{R}^n \Rightarrow \partial f(0) = \{s \in \mathbb{R}^n \mid \|y\| \geq s \cdot y\ \forall y \in \mathbb{R}^n\} = \{s \in \mathbb{R}^n \mid \max_{\|y\|=1} s \cdot y \leq 1\} = B_{\|\cdot\|_*}(0,1)$
- $f(x) = \|x\|_\infty,\ x \in \mathbb{R}^n \Rightarrow \partial f(0) = B_{\|\cdot\|_1}(0,1) = \operatorname{conv}(\pm e_1, \ldots, \pm e_n)$, $e_i = (0, \ldots, 1, 0, \ldots)$

Thm: Let $f_1, \ldots, f_m \in \operatorname{Conv} \mathbb{R}^n$, $t_1, t_2 > 0$, then

(a) $\partial(t_1 f_1 + t_2 f_2)(x) = t_1 \partial f_1(x) + t_2 \partial f_2(x)$ $\forall x \in \overbrace{\operatorname{int}(\operatorname{dom} f_1 \cap \operatorname{dom} f_2)}^{\text{necessary! (cf. } f_1 = I_{\{y \geq x^2\}},\ f_2 = I_{\{y \leq 0\}})}$ ("Moreau-Rockafellar")

(b) $\partial(\max_{i=1,\ldots,m} f_i)(x) = \operatorname{conv}\left(\bigcup_{j \in J_x} \partial f_j(x)\right)$ with $J_x = \{j \mid f_j(x) = f(x)\}$

proof: (b) without proof (see e.g. Bauschke & Combettes)

(a) "$\supseteq$": Let $s = t_1 s_1 + t_2 s_2$ with $s_i \in \partial f_i(x)$, $i = 1, 2$.

$$t_1 f_1(y) + t_2 f_2(y) \geq t_1 (f_1(x) + s_1 \cdot (y-x)) + t_2 (f_2(x) + s_2 \cdot (y-x))$$
$$= (t_1 f_1 + t_2 f_2)(x) + s \cdot (y-x)$$

Examples and properties cont'd

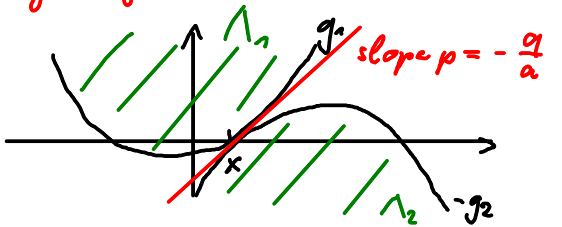
"$\subset$": Let $s \in \partial(t_1 f_1 + t_2 f_2)(x)$, set $g_1 = t_1 f_1 - t_1 f_1(x) - s\cdot(\cdot - x)$, $g_2 = t_2 f_2 - t_2 f_2(x)$

$\implies 0 \in \partial(g_1 + g_2)(x)$, $0 = g_1(x) = g_2(x)$, $\begin{cases} \partial g_2(x) = \partial t_2 f_2(x) = t_2 \partial f_2(x) \\ \partial g_1(x) = \partial t_1 f_1(x) - s = t_1 \partial f_1(x) - s \end{cases}$ why?

$\implies$ to show: $0 \in \partial g_1 + \partial g_2$

- $\Lambda_1 = \operatorname{epi} g_1$, $\Lambda_2 = -\operatorname{epi} g_2$ are convex & have nonempty interior (why?)
- $\operatorname{int} \Lambda_1 \cap \Lambda_2 = \emptyset$ (since $g_1 + g_2 \geq 0$ and thus $g_1 \geq -g_2$)

$\implies \exists\, 0 \neq \binom{q}{a} \in \mathbb{R}^n \times \mathbb{R}$: $\sup\limits_{(y,r)\in\Lambda_1} q\cdot y + ar \leq \inf\limits_{(y,r)\in\Lambda_2} q\cdot y + ar$



- due to $(x,0) \in \Lambda_1 \cap \Lambda_2$ even $\max\limits_{(y,r)\in\Lambda_1} q\cdot y + ar = q\cdot x = \min\limits_{(y,r)\in\Lambda_2} q\cdot y + ar$
- $a < 0$: $(x,r) \in \Lambda_1 \;\forall r > 0 \implies a \leq 0$, and $a = 0$ would imply $\sup\limits_{y\in\operatorname{dom} f_1} q\cdot y \leq \inf\limits_{y\in\operatorname{dom} f_2} q\cdot y$

which contradicts $\operatorname{int}(\operatorname{dom} f_1 \cap \operatorname{dom} f_2) \neq \emptyset$

- set $p = -\frac{q}{a}$ $\implies \underbrace{\max\limits_{(y,r)\in\Lambda_1} p\cdot y - r = p\cdot x}_{\text{choose } r = g_1(y) \implies p \in \partial g_1(x)} \overbrace{= \min\limits_{(y,r)\in\Lambda_2} p\cdot y - r}_{}$

choose $r = -g_2(y) \implies -p \in \partial g_2(x)$ $\square$

Subdifferential and conjugate

<u>Thm:</u> Let $f \in \operatorname{Conv} \mathbb{R}^n$. $s \in \partial f(x) \iff f^*(s) + f(x) = s \cdot x$

proof: $\cdot\, s \in \partial f(x) \iff s \cdot y - f(y) \leq s \cdot x - f(x) \;\; \forall y \in \operatorname{dom} f$

$\iff f^*(s) \leq s \cdot x - f(x)$

$\cdot\; f^*(s) \geq s \cdot x - f(x)$ by Fenchel's inequality

<u>Cor:</u> Let $f \in \operatorname{Conv} \mathbb{R}^n$ be closed. $s \in \partial f(x) \iff x \in \partial f^*(s)$

proof: $s \in \partial f(x) \iff f^*(s) + f(x) = s \cdot x \iff f^*(s) + f^{**}(x) = s \cdot X \iff x \in \partial f^*(s)$ $\square$

<u>Cor:</u> Let $f \in \operatorname{Conv} \mathbb{R}^n$ be closed, then

$$x^* = \operatorname*{argmin}_{x \in \mathbb{R}^n} f(x) \iff 0 \in \partial f(x^*) \iff x^* \in \partial f^*(0)$$

In particular, if f^* is differentiable in 0, the minimiser is unique.

Slater's constraint qualification

Thm: (Slater) Consider $\min_{x \in \mathbb{R}^n} f_0(x)$ s.t. $f_1(x), \dots, f_m(x) \le 0$, $h_1(x) = \dots = h_p(x) = 0$ (P)

for $f_0, \dots, f_m$ convex, $h_1, \dots, h_p$ affine (i.e. $h_1(x) = \dots = h_p(x) = 0 \iff Ax = b$), $\mathcal{D} = \bigcap_{i=0}^{m} \operatorname{dom} f_i$.

If (P) is strictly feasible, i.e. $\exists\, \tilde{x} \in \operatorname{relint} \mathcal{D}$ with $f_i(\tilde{x}) < 0$, $i = 1, \dots, m$, $A\tilde{x} = b$,

then strong duality holds, i.e. $p^* = d^*$.

for proof without see later!

proof: · Simplifying assumptions: $\operatorname{relint} \mathcal{D} = \operatorname{int} \mathcal{D}$, $\operatorname{rank} A = p$

(both can be achieved by transformation to equivalent problems: cancel linearly dependent rows of $Ax = b$ & restrict to $\operatorname{aff} \mathcal{D}$.

Note: dual problem might change by this!)

· $\tilde{x}$ is feasible $\Rightarrow p^* < \infty$. Wlog $p^* > -\infty$ (else $d^* = -\infty$ by weak duality)

· define $M = \{(u, v, t) \in \mathbb{R}^{m+p+1} \mid \exists x \in \mathcal{D} : f_i(x) \le u_i, i = 1, \dots, m, h_i(x) = v_i, i = 1, \dots, p, f_0(x) \le t\}$

$N = \{(0, 0, s) \in \mathbb{R}^{m+p+1} \mid s < p^*\}$

$\Rightarrow M, N$ convex & $M \cap N = \emptyset$ (otherwise there were some feasible x with $f_0(x) < p^*$)

Slater's constraint qualification cont'd

- there exists a separating hyperplane, i.e. $(\mu, \lambda, \nu) \neq 0$, $\alpha \in \mathbb{R}$ s.t.

$$\text{(a)} \quad \mu \cdot u + \lambda \cdot v + \nu t \geq \alpha \qquad \forall (u,v,t) \in M$$

$$\text{(b)} \quad \mu \cdot u + \lambda \cdot v + \nu t \leq \alpha \qquad \forall (u,v,t) \in N$$

- (a) $\Rightarrow \mu \geq 0$ (comp.-wise), $\nu \geq 0$ (else $\mu \cdot u + \nu t$ would be unbdd below in M)

  (b) $\Rightarrow \nu p^* \leq \alpha$

- together, $\sum_{i=1}^m \mu_i f_i(x) + \sum_{i=1}^p \lambda_i h_i(x) + \nu f_0(x) \geq \alpha \geq \nu p^* \quad \forall x \in \mathcal{D}$

  case $\nu > 0$: $\Rightarrow L(x, \mu/\nu, \lambda/\nu) \geq p^* \Rightarrow g(\mu/\nu, \lambda/\nu) \geq p^* \Rightarrow d^* \geq p^*$

  case $\nu = 0$: $\Rightarrow \sum_{i=1}^m \mu_i f_i(\tilde{x}) \geq 0 \underset{\mu_i \geq 0,\ f_i(\tilde{x}) < 0}{\Longrightarrow} \mu = 0$

  $\Rightarrow \sum_{i=1}^p \lambda_i h_i(x) \geq 0 \ \forall x \in \mathcal{D}$

  however, $\sum_{i=1}^p \lambda_i h_i(\tilde{x}) = 0$ together with rank $A = p$

  implies existence of $x \in \mathcal{D}$ with $\sum_{i=1}^p \lambda_i h_i(x) < 0$ ↯

  $\Rightarrow \nu = 0$ impossible $\square$

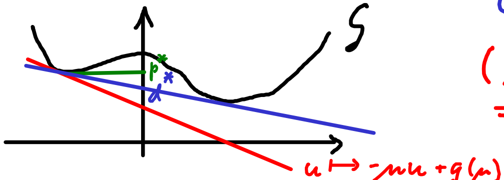
<u>Cor:</u> $(\mu^*, \lambda^*) = (\mu/\nu, \lambda/\nu)$ are dual optimal!

Slater's constraint qualification - geometric intuition

$$\mathcal{G} = \{(u,v,t) \in \mathbb{R}^{m+p+1} \mid \exists x \in \mathcal{D},\ f_i(x) = u_i,\ h_i(x) = v_i,\ f_0(x) = t\}$$

recall: $g(\mu,\lambda) = \inf_x f_0(x) + \sum_i \mu_i f_i(x) + \sum_i \lambda_i h_i(x) = \inf_{(u,v,t)\in\mathcal{G}} t + \mu^T u + \lambda^T v$; $\mu \geq 0$

Ex: $\min_x f_0(x)$ s.t. $f_1(x) \leq 0$

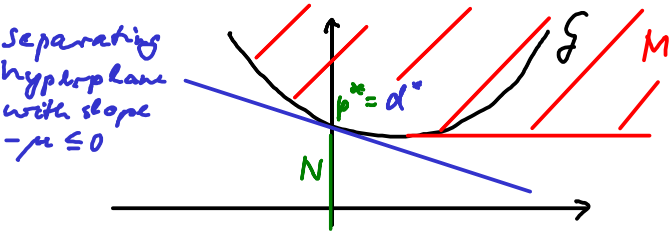


(let $\ell_c(u,v) = -\mu^T u - \lambda^T v + c$

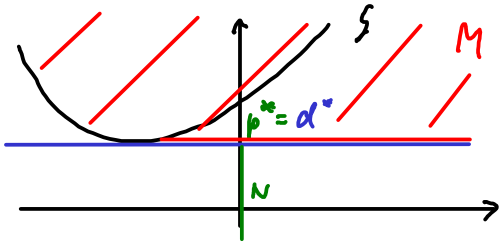
$\Rightarrow \inf\{\ell_c(0) \mid \exists (u,v) : (u, v, \ell_c(u,v)) \in \mathcal{G}\}$

$= \inf\{t + \mu^T u + \lambda^T v \mid (u,v,t) \in \mathcal{G}\}$)

Slater's condition holds

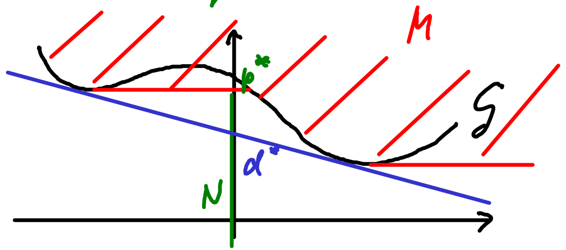


Slater's condition holds

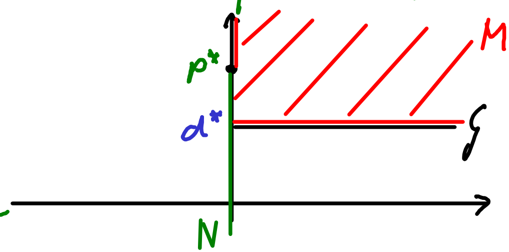


M & N cannot be separated by hyperplane with finite slope

nonconvex problem



no strictly feasible point



Implications of constraint qualification

Thm: Strong duality even holds if there is a feasible $\tilde{x} \in \operatorname{relint} \mathcal{D}$ with $f_i(\tilde{x}) < 0$ only for all nonaffine f_i.

proof: see Rockafellar: Convex Analysis, Thm 28.2 $\square$

In particular, we always have strong duality if all constraints are affine and f_0 is convex, thus for all LPs and QPs!

Thm: Under a constraint qualification (i.e. a condition implying strong duality such as Slater's), if the dual problem is feasible, the KKT conditions hold for a convex optimisation problem (P) at an optimal point x, i.e. $\exists \mu, \lambda$ s.t.

$$f_1(x), \dots, f_m(x) \le 0, \quad h_1(x), \dots, h_p(x) = 0, \quad \mu \ge 0, \quad \mu^T \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} = 0,$$

$$0 \in \partial \Big[f_0 + \sum_{i=1}^m \mu_i f_i + \sum_{i=1}^p \lambda_i h_i \Big](x)$$

proof: strong duality $\Rightarrow$ saddle point property of L $\Longrightarrow$ x minimizes $L(\cdot, \mu, \lambda)$ $\square$

take μ, λ dual optimal (possible, already shown)

Fenchel-Rockafellar-duality

Thm: (Fenchel-Rockafellar) Let $f, g \in \operatorname{Conv} \mathbb{R}^n$. If either

(a) $\operatorname{relint} \operatorname{dom} f \cap \operatorname{relint} \operatorname{dom} g \neq \emptyset$ or

(b) $\operatorname{relint} \operatorname{dom} f^* \cap -\operatorname{relint} \operatorname{dom} g^* \neq \emptyset$ and f, g closed,

then $\inf_{x \in \mathbb{R}^n} f(x) + g(x) = \sup_{y \in \mathbb{R}^n} -g^*(-y) - f^*(y)$.

Under (a) the supremum is attained, under (b) the infimum.

proof: · $\forall x, y \in \mathbb{R}^n$: $f(x) + f^*(y) \geq x \cdot y \geq -g(x) - g^*(-y)$ (Fenchel ineq.)

$\Rightarrow \inf_x (f(x) + g(x)) \geq \sup_y -g^*(-y) - f^*(y)$

· if $\inf = -\infty$, also $\sup = -\infty$, thus assume wlog $-\infty < \alpha = \inf_x f(x) + g(x)$

· Let (a) hold; it suffices to show existence of some $y \in \mathbb{R}^n$ with $-g^*(-y) - f^*(y) \geq \alpha$

· Set $C = \{(x, \nu) \in \mathbb{R}^n \times \mathbb{R} \mid \nu \geq f(x)\}$

$D = \{(x, \nu) \in \mathbb{R}^n \times \mathbb{R} \mid \nu \leq -g(x) + \alpha\}$

$\Rightarrow C, D$ convex & $\operatorname{relint} C = \{(x, \nu) \mid x \in \operatorname{relint} \operatorname{dom} f, f(x) < \nu\}$ is disjoint from D

Fenchel-Rockafellar-duality cont'd

- $\exists$ separating hyperplane, i.e. $(s,t) \in \mathbb{R}^n \times \mathbb{R}$ with $\sup_{(x,\nu)\in D} s\cdot x + t\nu \le \inf_{(x,\nu)\in C} s\cdot x + t\nu$.

  Note: $t \neq 0$ (otherwise the projections of D & C, dom f & dom g, would be separated)

  $\Rightarrow$ separating hyperplane is graph of an affine function $h(x) = \underset{-s/t}{\underset{\shortparallel}{\tilde{s}}} \cdot x + \beta$

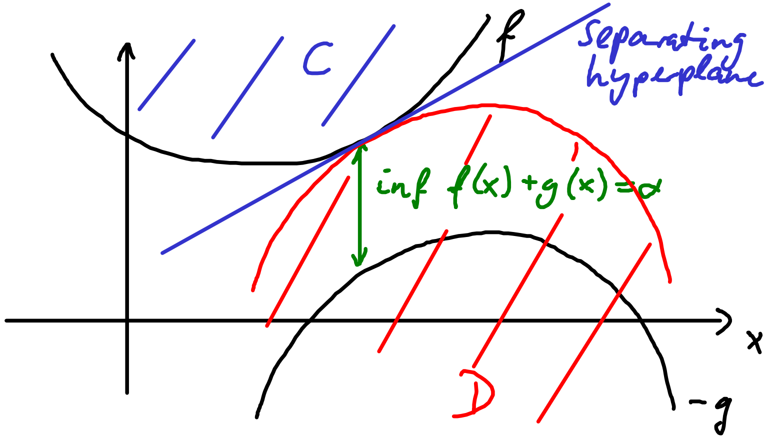
  $\Rightarrow f(x) \ge \tilde{s}\cdot x + \beta \ge -g(x) + \alpha \quad \forall x \in \mathbb{R}^n$

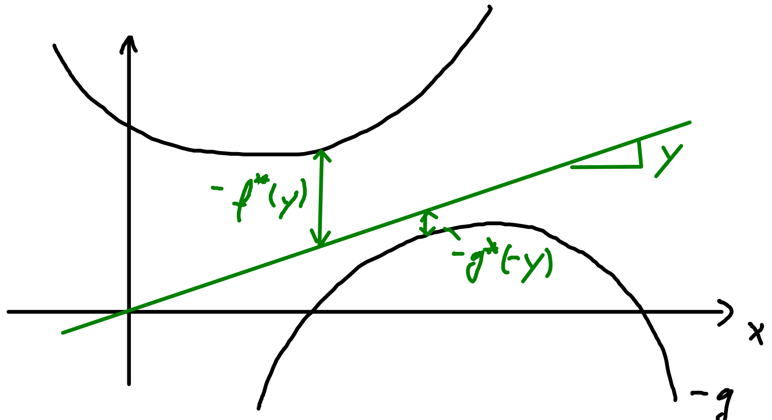
- left inequality implies $-\beta \ge \sup_x \tilde{s}\cdot x - f(x) = f^*(\tilde{s})$

  right ineq. implies $\alpha - \beta \le \inf_x \tilde{s}\cdot x + g(x) = -g^*(-\tilde{s})$

  $\Bigg\} \Rightarrow \alpha \le -g^*(-\tilde{s}) - f^*(\tilde{s})$

- for ($\&$), all follows from duality, noting $f = f^{**}$, $g = g^{**}$. $\square$





Fenchel-Rockafellar-duality: extensions

The result can be sharpened & generalized in different ways, e.g.

- if g is piecewise affine, we may replace relint dom $g^{(*)}$ by dom $g^{(*)}$

(Rockafellar, Thm. 31.1)

- if $A:\mathbb{R}^n \to \mathbb{R}^m$ is a linear operator, $\inf_{x\in\mathbb{R}^n} f(x)+g(Ax) = \sup_{y\in\mathbb{R}^m} -g^*(-y)-f^*(A^Ty)$ if

  (a) $\exists\, x \in$ relint dom f with $Ax \in$ relint dom g or

  (b) $\exists\, y \in$ relint dom g^* with $-A^Ty \in$ relint dom f^*

(Rockafellar, Cor. 31.2.1)

- under the same conditions, $\inf_{x\in\mathbb{R}^n} (f(x)+g(Ax)) = \inf_{x\in\mathbb{R}^n} \sup_{y\in\mathbb{R}^m} x^TA^Ty + f(x) - g^*(y)$ and inf & sup may be swapped.

proof: $\inf_{x\in\mathbb{R}^n} (f(x)+g(Ax)) \geq \inf_{x\in\mathbb{R}^n} (f(x)+g^{**}(Ax)) = \inf_{x\in\mathbb{R}^n} \sup_{y\in\mathbb{R}^m} x^TA^Ty + f(x) - g^*(y)$

$\geq \sup_{y\in\mathbb{R}^m} \inf_{x\in\mathbb{R}^n} x^TA^Ty + f(x) - g^*(y) = \sup_{y\in\mathbb{R}^m} -f^*(-A^Ty) - g^*(y) = \sup_{y\in\mathbb{R}^m} -f^*(A^Ty) - g^*(-y)$

$\Rightarrow$ strong duality implies equality $\square$

Fenchel-Rockafellar-duality & Lagrange-duality

Rem: For convex f_0, c_I, c_E, Lagrange duality is a special case of Fenchel-Rockafellar duality:

choose $f((x,u,v)) = f_0(x) + I_{\{c_I(x) \le u,\ c_E(x) = v\}}((x,u,v))$, $h((x,u,v)) = I_{(-\infty,0]^m \times \{0\}}((u,v))$

$\Rightarrow f^*((y,\mu,\lambda)) = \sup\limits_{x,u,v,\ u \ge c_I(x),\ v = c_E(x)} y\cdot x + \mu\cdot u + \lambda\cdot v - f_0(x)$, $h^*((y,\mu,\lambda)) = I_{\{0\}\times[0,\infty)^m}((y,\mu$

Now Slater $\Rightarrow \exists\, x \in \operatorname{relint} \mathcal{D}$ with $f((x,-\varepsilon,0)) < \infty$ for $\varepsilon > 0$ small enough (comp.-wise)

$$\Rightarrow \underbrace{\operatorname{relint} \operatorname{dom} h}_{\mathbb{R}^n \times (-\infty,0)^m \times \{0\}} \cap \operatorname{relint} \operatorname{dom} f \neq \emptyset$$

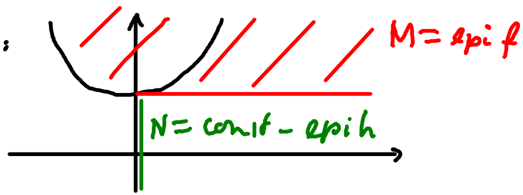
$$\Rightarrow p^* = \inf_{x,\ c_I(x) \le 0,\ c_E(x) = 0} f_0(x) = \inf_{(x,u,v)} f((x,u,v)) + h((x,u,v)) \overset{\text{Fenchel-Rockafellar}}{=} \max_{(y,\tilde\mu,\tilde\lambda)} -f^*((y,\tilde\mu,\tilde\lambda)) - h^*(-(y,\tilde\mu,\tilde\lambda))$$

$$= \max_{\tilde\mu \le 0,\tilde\lambda} \inf_{\substack{x,u,v \\ u \ge c_I(x),\ v = c_E(x)}} -\tilde\mu\cdot u - \tilde\lambda\cdot v + f_0(x) = \max_{\tilde\mu,\tilde\lambda} \begin{Bmatrix} \inf_x -\tilde\mu\cdot c_I(x) - \tilde\lambda\cdot c_E(x) + f_0(x) & \text{if } \tilde\mu \le 0 \\ \infty & \text{else} \end{Bmatrix} = \max_{\mu \ge 0,\lambda} g(\mu,\lambda)$$

$$= d^*$$

This also implies existence of a dual optimal (μ,λ). $\square$

recall:



Linear program basics

how to obtain standard form? hint: • $x = x_+ - x_-$ • new variable for ineq. constraints

Def: • A linear program (LP) in standard form is given by

$$\min_{x \in \mathbb{R}^n} cx \quad \text{s.t.} \quad Ax = b, \quad x \geq 0 \text{ coordinate-wise,} \qquad (LP)$$

where $c \in \mathbb{R}^{1\times n}$, $A \in \mathbb{R}^{p \times n}$ has full rank, $p < n$.

It can be viewed as minimisation of cx over the convex polyhedron $K = \{x \geq 0 \mid Ax = b\}$.

• Let $B \in \mathbb{R}^{p\times p}$ be a matrix composed of m linearly independent columns a_i, $i \in J \subset \{1, \dots, n\}$ of A. A point $x \in \mathbb{R}^n$ with $x_i = 0 \; \forall i \notin J$ and $Ax = b$ is called a basic point wrt. the basis B.

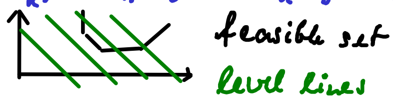
• A feasible point x satisfies $Ax = b$, $x \geq 0$; a basic feasible point is basic & feasible. We shall assume that any basic feasible point has $x_i \neq 0 \; \forall i \in J$.

Thm: If (LP) admits optimal points, then at least one of them is an extreme point of K.

proof: Let $\operatorname{ext} K = \{x_1, \dots, x_k\}$, then $\forall x \in K \; \exists \alpha_1, \dots, \alpha_k \geq 0 : \alpha_1 + \dots + \alpha_k = 1$, $x = \alpha_1 x_1 + \dots + \alpha_k x_k$.

Let $z_0 = \min_{i=1,\dots,k} cx_i$, then $cx = \alpha_1(cx_1) + \dots + \alpha_k(cx_k) \geq \alpha_1 z_0 + \dots + \alpha_k z_0 = z_0 \Rightarrow z_0$ is min! $\square$

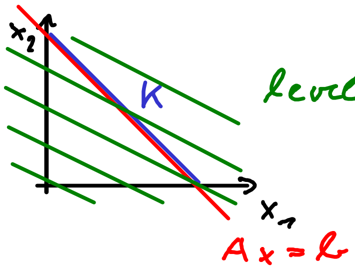
note: This does not depend on normal form!



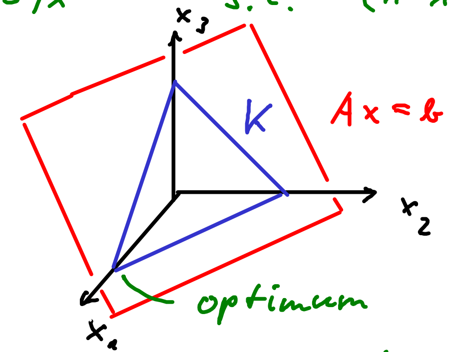
Examples

<u>Ex:</u>

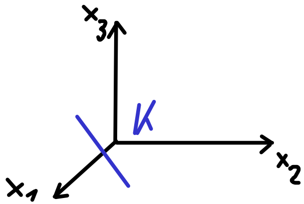
- $\min\limits_{x\in\mathbb{R}^2} \underbrace{(1\ \ 2)}_{c}\, x$ s.t. $\overbrace{(1\ \ 1)}^{A}\, x = \underset{\sim}{\overset{b}{2}}$, $x \geq 0$



- $\min\limits_{x\in\mathbb{R}^3} (1\ 2\ 3)\,x$ s.t. $(1\ 1\ 1)\,x = 3$, $x \geq 0$



- $\min\limits_{x\in\mathbb{R}^3} (1\ 2\ 3)\,x$ s.t. $\begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix} x = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$, $x \geq 0$



note: extreme points of K have n - rank A zeros!
$\Longrightarrow$ basic points ?

optimum ?

Basic solutions & extreme points

Thm: $x \in \operatorname{ext} K \iff x$ is a basic feasible point

proof: "$\Rightarrow$" • wlog assume the nonzero entries of x to be $x_1, \dots, x_k$

$\Rightarrow b = x_1 a_1 + \dots + x_k a_k$ with $x_i > 0$ (columns of A)

• assume $a_1, \dots, a_k$ are linearly dependent, i.e. $0 = y_1 a_1 + \dots + y_k a_k$ for some y_i

• letting $y = (y_1, \dots, y_k, 0, \dots, 0)$, there exists $\varepsilon > 0$ with $x + \varepsilon y \geq 0$, $x - \varepsilon y \geq 0$

$\Rightarrow x = \frac{1}{2} \underbrace{(x + \varepsilon y)}_{\in K} + \frac{1}{2} \underbrace{(x - \varepsilon y)}_{\in K}$, which contradicts $x \in \operatorname{ext} K$

"$\Leftarrow$" • wlog assume $x = (x_1, \dots, x_p, 0, \dots, 0)$ (i.e. $a_1, \dots, a_p$ form a basis)

• assume there are $y, z \in K$, $\alpha \in (0,1)$ with $x = \alpha y + (1-\alpha) z$

$\Rightarrow y_{p+1} = \dots = y_n = 0$ & $z_{p+1} = \dots = z_n = 0$ (due to $y, z \geq 0$)

$\Rightarrow y_1 a_1 + \dots + y_p a_p = b$ & $z_1 a_1 + \dots + z_p a_p = b$

Since $a_1, \dots, a_p$ are linearly independent, this implies $y = z$

$\Rightarrow x \in \operatorname{ext} K$ $\square$

Simplex method: the idea

$\#\,\text{ext}\,K$ = # basic feasible points $\leq$ # possibilities to choose p lin. indep. columns from A

↑ number of...

$$\leq \binom{n}{p} = \frac{n!}{p!\,(n-p)!}$$

Hence, to find an optimal point one only has to test at most $\binom{n}{p}$ extreme points.

The simplex algorithm does better by generating a sequence x^k of extreme points, where the function value improves in each step.

Algorithm (Simplex method):

0) (Initialisation) pick a basis / an extreme point

- choose columns $a_{i_1}, \dots, a_{i_p}$, $i_j \in \mathcal{J}^0$, of A as a basis
- set $x^0_\ell = 0$ for $\ell \notin \mathcal{J}^0$ and solve $Ax^0 = b \iff a_{i_1} x^0_{i_1} + \dots + a_{i_p} x^0_{i_p} = b$ for $x^0_{i_1}, \dots, x^0_{i_p}$
- if the basic point is infeasible, choose a different basis
- note: we may transform $Ax = b$ via Gaussian elimination s.t. $a_{i_j} = e_j$
- note: superscript = iteration ; subscript = column / row index

Simplex method: the algorithm

iterate:

1) choose function-decreasing direction

- let x^{k+1} be the next iterate to be found
- $x_{i_1}^{k+1}, \dots, x_{i_p}^{k+1}$ can be solved for in terms of $x_\ell^{k+1}, \ell \notin J^k$: $x_{i_j}^{k+1} = \alpha_{j0} + \sum_{\ell \notin J^k} \alpha_{j\ell} x_\ell^{k+1}$ ← easy if $a_{i_j} = e_j$
- $c x^{k+1} = z_0 + \sum_{\ell \notin J^k} (c_\ell + z_\ell) x_\ell^{k+1}$ for $z_\ell = \sum_{j=1}^{p} c_{i_j} \alpha_{j\ell}$
- pick $\ell \notin J^k$ with $c_\ell + z_\ell < 0$, then increasing x_ℓ from 0 will decrease cx

  if there is no such ℓ, x^k is optimal

2) move to an adjacent basis / extreme point

- want to add a_ℓ to basis; which of $a_{i_1}, \dots, a_{i_p}$ should it replace?
- $a_\ell = y_1 a_{i_1} + \dots + y_p a_{i_p} \Rightarrow (x_{i_1}^k - \alpha y_1) a_{i_1} + \dots + (x_{i_p}^k - \alpha y_p) a_{i_p} + \alpha a_\ell = b \quad \forall \alpha \in \mathbb{R}$
- choose $x_{i_j}^{k+1} = x_{i_j}^k - \alpha y_j$, $j = 1, \dots, p$, $x_\ell^{k+1} = \alpha$ for some α

if $y_1, \dots, y_p \leq 0$: x^{k+1} cannot become a basic feasible point & $p^* = -\infty$

else: let $q = \operatorname{argmin}_{s=1,\dots,p,\, y_s > 0} \frac{x_{i_s}^k}{y_s}$; set $\alpha = \frac{x_{i_q}^k}{y_q}$, $i_q = \ell$, $J^{k+1} = \{i_1, \dots, i_p\}$

may again transform $Ax = b$ s.t. $a_{i_j} = e_j$! the only choice to make x^{k+1} basic feasible

Simplex method examples

Ex: $\min\limits_{x\in\mathbb{R}^3} \;(1\ \ 2\ \ 3)\,x \quad$ s.t. $\quad (1\ \ 1\ \ 1)\,x = 3\,,\quad x \geq 0$

0) $\mathcal{J}^0 = \{3\} \Longrightarrow x^0 = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} \geq 0$

1) $(1\ \ 1\ \ 1)\,x = 3 \Longrightarrow x_3^1 = 3 - x_1^1 - x_2^1 \Longrightarrow (1\ \ 2\ \ 3)\,x^1 = x_1^1 + 2x_2^1 + 3(3 - x_1^1 - x_2^1)$
$= 9 \underbrace{-\,2}_{c_1 - z_1} x_1^1 \underbrace{-\,1}_{c_2 - z_2} x_2^1 \Longrightarrow$ pick $l = 2$!

2) trivial: $\mathcal{J}^1 = \{2\} \Longrightarrow x^1 = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}$

1) $x_2^2 = 3 - x_1^2 - x_3^2 \Longrightarrow c\,x^2 = 6 - x_1^2 + x_3^2 \Longrightarrow$ pick $l = 1$!

2) trivial: $\mathcal{J}^2 = \{1\} \Longrightarrow x^2 = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix}$

1) $x_1^3 = 3 - x_2^3 - x_3^3 \Longrightarrow c\,x^3 = 3 + x_2^3 + 2x_3^3 \Longrightarrow x^3$ is optimal!

Ex: $\min\limits_{x\in\mathbb{R}^3} \;(1\ \ 2\ \ 3)\,x \quad$ s.t. $\quad \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix} x = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \quad x \geq 0$

0) $\mathcal{J}^0 = \{1, 3\} \Longrightarrow x^0 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$

1) $Ax = b \Longrightarrow \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x_1^1 \\ x_3^1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} x_2^1 \Longrightarrow \begin{pmatrix} x_1^1 \\ x_3^1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 - x_2^1 \end{pmatrix}$

$\Longrightarrow c\,x = 1 + 2x_2^1 + 3(2 - x_2^1) \Longrightarrow$ pick $l = 2$!

2) $a_2 = 0 \cdot a_1 + 1 \cdot a_3 \Longrightarrow (x_1^0 - \alpha \cdot 0)\,a_1 + (x_3^0 - \alpha \cdot 1)\,a_3 + \alpha\, a_2 = b$

$\Longrightarrow q = 3,\ \alpha = x_3^0 = 2 \Longrightarrow \mathcal{J} = \{1, 2\},\ x^2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$

1) $\Longrightarrow x^2$ is optimal!

Structure of Linesearch methods

In the following we shall assume $f : \mathbb{R}^n \to \mathbb{R}$ to be twice differentiable and bounded from below.

<u>Def</u>: A <u>descent direction</u> for f at x is a $p \in \mathbb{R}^n$ with $Df(x)p < 0$
(direction in which f decreases)

<u>Alg. (Linesearch method)</u>:

given: $f : \mathbb{R}^n \to \mathbb{R}$, $x_0 \in \mathbb{R}^n$

repeat
1) choose a descent direction $p_k \in \mathbb{R}^n$
2) choose a step length $\alpha_k > 0$
3) $x_{k+1} = x_k + \alpha_k p_k$
4) $k \leftarrow k+1$

until x_{k+1} sufficiently minimises f

Stepsize control: Armijo condition

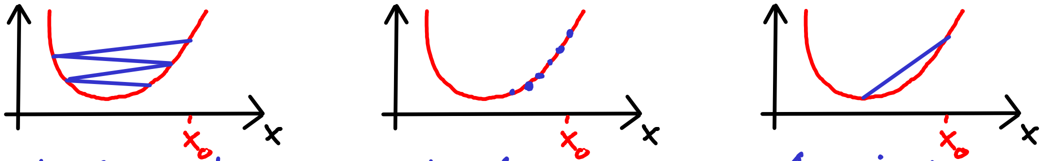
Def: $\alpha_k > 0$ is said to satisfy **Armijo's condition** for $0 < c_1 < 1$ if

$$f(x_k + \alpha_k p_k) \le f(x_k) + c_1 \alpha_k Df(x_k) p_k$$

expected descent!

Cor: If $f \in C^1(\mathbb{R}^n)$, $0 < c_1 < 1$, there exists a step length α satisfying Armijo's condition

proof: Taylor's thm $\Rightarrow$ $f(x + \alpha p) = f(x) + \alpha Df(x) p + o(\alpha) < f(x) + c_1 \alpha Df(x) p$ for α small enough $\square$



too long steps — too short steps — Armijo step

Alg. (backtracking to find good step length):

$\alpha = 1$

if Armijo condition fulfilled

  repeat $\alpha \leftarrow 2\alpha$ until Armijo condition violated

while Armijo condition violated

  $\alpha \leftarrow \alpha / 2$

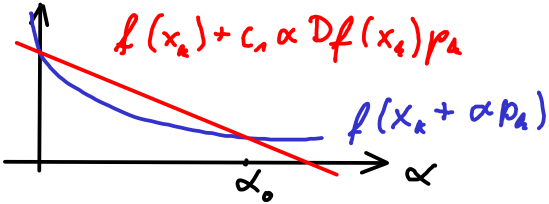
Stepsize control: Wolfe conditions

Def: α_k is said to satisfy the strong Wolfe conditions for $0<c_1<c_2<1$ if it satisfies Armijo's cond. &

$$|Df(x_k+\alpha_k p_k)p_k| \leq -c_2\, Df(x_k)p_k$$

Thm: If $f\in C^1(\mathbb{R}^n)$ bounded from below, $0<c_1<c_2<1$, there exists α_k satisfying the strong Wolfe cond.

proof: • Let $g(\alpha) = f(x_k+\alpha p_k)$



• $\alpha_0 := \max\{\alpha\in\mathbb{R} \mid f(x_k+\alpha p_k)\leq f(x_k)+c_1\alpha Df(x_k)p_k\}$

• if $g'(\alpha_0)\geq 0$ choose $\alpha_k = \max\{\alpha<\alpha_0 \mid g'(\alpha)=0\}$

note: automatically satisfies Armijo's condition

• else $|Df(x_k+\alpha_0 p_k)p_k| = -g'(\alpha_0) < -\underbrace{c_1 g'(0)}_{\text{slope of red line}} < -c_2\, g'(0) = -c_2 Df(x_k)p_k$ $\square$

Global convergence

Thm (Zoutendijk): If
- $\exists L>0: \|\nabla f(x)-\nabla f(y)\|_2 \le L\|x-y\|_2$, (gradient is Lipschitz)
- $\cos\theta_k = \dfrac{-\nabla f(x_k)p_k}{\|\nabla f(x_k)\|_2 \|p_k\|_2}$, (angle between $-\nabla f(x_k)$ and search direction)
- α_k satisfies strong Wolfe conditions,

then $\sum_{k=1}^{\infty} \cos^2\theta_k \|\nabla f(x_k)\|_2^2 < \infty$ (either $\nabla f \to 0$ or angle degenerate)

proof: 1) use Wolfe condition & fact that ∇f cannot change too fast to obtain lower bound on step length

$$\alpha_k L \|p_k\|_2^2 \ge (\nabla f(x_k+\alpha_k p_k) - \nabla f(x_k))p_k \ge (c_2-1)\nabla f(x_k)p_k$$

$$\Rightarrow \alpha_k \ge \frac{c_2-1}{L\|p_k\|_2^2}\nabla f(x_k)p_k$$

2) use Armijo condition to estimate worst case function decrease

$$\Rightarrow f(x_{k+1}) \le f(x_k) - c_1\frac{1-c_2}{L}\frac{(\nabla f(x_k)p_k)^2}{\|p_k\|_2^2} = f(x_k) - c_1\frac{1-c_2}{L}\cos^2\theta_k\|\nabla f(x_k)\|_2^2$$

the subsequent convergence rate proofs essentially start from here

$$\Rightarrow -\infty < \inf f - f(x_0) \le f(x_\ell) - f(x_0) = \sum_{k=0}^{\ell-1} f(x_{k+1}) - f(x_k) \le -c_1\frac{1-c_2}{L}\sum_{k=0}^{\ell-1}\cos^2\theta_k\|\nabla f(x_k)\|_2^2 \quad \square$$

Rem: If ∇f is only Hölder with exponent β, one gets $\sum_{k=1}^{\infty} \left|\cos\theta_k \|\nabla f(x_k)\|_2\right|^{1+\frac{1}{\beta}}$ (homework)

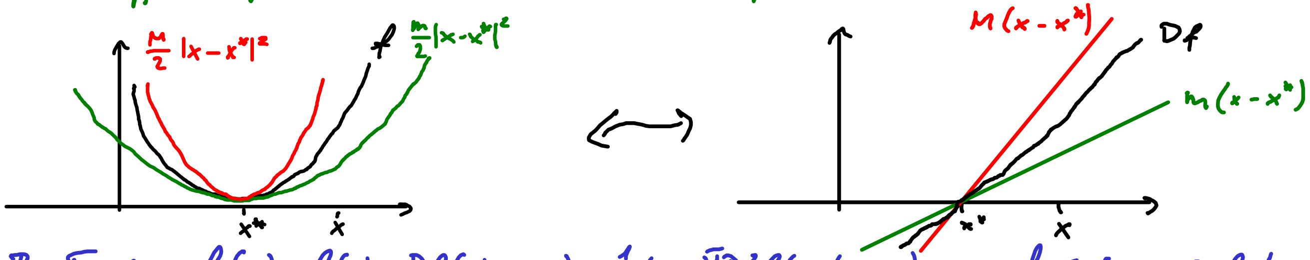
$\Rightarrow \nabla f \to 0$ at a slower rate!

Optimality bounds from gradient

Thm: If there exist $m, M > 0$ s.t. $\underbrace{mI \le D^2 f(x) \le MI}_{\text{in sense of quadratic forms, i.e. eigenvalues lie in } [m,M]}$ $\forall\, x \in \mathcal{U}(x^*)$, then $\forall\, x \in \mathcal{U}(x^*)$

neighbourhood of opt. point x^*

- $\frac{1}{2M} \|Df(x)\|_2^2 \le f(x) - p^* \le \frac{1}{2m} \|Df(x)\|_2^2$
- $\frac{1}{M} \|Df(x)\|_2 \le \|x - x^*\|_2 \le \frac{1}{m} \|Df(x)\|_2$



proof: • By Taylor, $f(y) = f(x) + Df(x)(y-x) + \frac{1}{2}(y-x)^T D^2 f(z)(x-y)$ for some z between x, y

$\Longrightarrow \quad f(y) \overset{\le}{\ge} f(x) + Df(x)(y-x) + \frac{\overset{M}{m}}{2} \|y-x\|_2^2 \quad \forall\, x, y \in \mathcal{U}$ (if $\mathcal{U}$ convex)

• minimise both sides over y: on lhs, $y = x^*$

on rhs, $y = \tilde{y} = x - \frac{1}{\overset{M}{m}} \nabla f(x)$

$\Longrightarrow \quad p^* = f(x^*) \overset{\le}{\ge} f(x) - \frac{1}{2\overset{M}{m}} \|Df(x)\|_2^2$

• As above, $f(x) \overset{\le}{\ge} f(x^*) + Df(x^*)(x-x^*) + \frac{\overset{M}{m}}{2} \|x-x^*\|_2^2 = p^* + \frac{\overset{M}{m}}{2} \|x - x^*\|_2^2$

$\Longrightarrow \quad \|x - x^*\|_2^2 \overset{\ge}{\le} \frac{2}{\overset{M}{m}} (f(x) - p^*) \overset{\ge}{\le} \frac{1}{\overset{M^2}{m^2}} \|Df(x)\|_2^2$ $\square$

Gradient descent, steepest descent, Newton's method

Def: • The gradient descent direction is $p_k^g = -\nabla f(x_k)$

• Let $\|\cdot\|$ be a norm with dual norm $\|\cdot\|_*$; the steepest descent direction is

$$p_k^s = \|\nabla f(x_k)\|_* \underset{\|v\|=1}{\operatorname{argmin}} Df(x_k)v$$

(gradient descent = steepest descent with $\|\cdot\|_2$)

(scaling s. t. $-Df(x_k)p_k^s = \|\nabla f(x_k)\|_*^2$)

• The Newton step is $p_k^N = -D^2 f(x_k)^{-1} \nabla f(x_k)$

Rem: If $D^2 f$ is positive definite, all above are descent directions. Indeed,

$$Df(x_k)p_k^g = -\|Df(x_k)\|_2^2 < 0, \quad Df(x_k)p_k^s = -\|\nabla f(x_k)\|_*^2 < 0, \quad Df(x_k)p_k^N = -\|\nabla f(x_k)\|_{D^2 f(x_k)}^2 < 0.$$

Rem: The Newton step finds minimum in a single step if f is quadratic and $\alpha_k = 1$:

$f(x) = \frac{1}{2}x^T A x + b^T x$ is minimised by $x^* = -A^{-1}b$, and $x_k + \alpha_k p_k^N = x_k - A^{-1}(Ax_k + b) = x^*$.

Cor: If $mI \le D^2 f(x) \le MI \ \forall x \in \mathbb{R}^n$, the α_k satisfy the strong Wolfe conditions, then

$$Df(x_k) \xrightarrow[k\to\infty]{} 0 \quad \text{for gradient/steepest/Newton descent.}$$

proof: Follows from Zoutendijk's theorem with $\cos\theta_k = 1$ for gradient descent,

$\cos\theta_k \ge \frac{m}{M}$ for Newton's method, $\cos\theta_k \ge \frac{c}{C}$ for steepest descent with $c\|\cdot\| \le \|\cdot\|_2 \le C\|\cdot\|_*$ $\square$

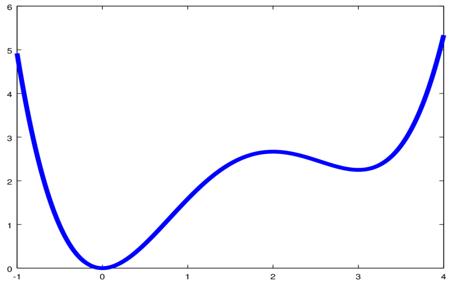
Experimental convergence rate

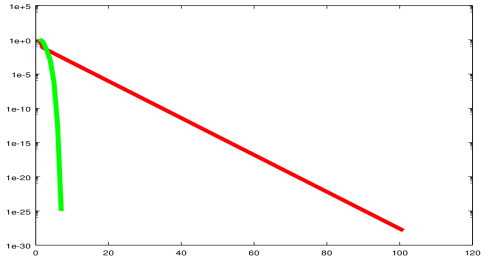
```
function experimentalConvergenceRate
  % f(x) = x^4/4-5x^3/3+3x^2
  % f'(x) = x*(x-2)*(x-3) = x^3-5x^2+6x; global minimum at 0
  f = @(x) x.^4/4-5*x.^3/3+3*x.^2;
  df = @(x) x*(x-2)*(x-3);
  d2f = @(x) 3*x^2-10*x+6;
  fun = @(x) combineFunctions(x,f,df,d2f);

  x = -1:.01:4;
  plot(x,f(x),'Linewidth',5);
  pause;

  x0 = 1.3;
  maxIter = 100;
  [x,iterSteepest] = descendSteepest( fun, x0, maxIter, true );
  [x,iterNewton] = newtonMethod( fun, x0, maxIter, 1e-26, true );
  semilogy(1:length(iterSteepest),abs(iterSteepest),'r','Linewidth',5,...
           1:length(iterNewton),abs(iterNewton),'g','Linewidth',5);
end

function [A,B,C] = combineFunctions(x,a,b,c)
  A = a(x);
  B = b(x);
  C = c(x);
end
```





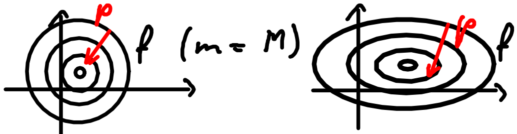
Convergence rate gradient/steepest descent

Thm: Let $S=\{x\in\mathbb{R}^n \mid f(x)\le f(x_0)\}$, $mI \le D^2 f(x) \le MI$ $\forall x\in S$, $\left\{\begin{matrix} \frac{1}{\gamma}\|\cdot\| \ge \|\cdot\|_2 \ge \gamma\|\cdot\|_* \\ \frac{1}{\tilde\gamma}\|\cdot\|_* \ge \|\cdot\|_2 \ge \tilde\gamma\|\cdot\| \end{matrix}\right\}$,

x_k the steepest descent iterates for backtracking with $c_1\in(0,\frac{1}{2})$. Then

$$(f(x_k)-p^*) \le \delta^k (f(x_0)-p^*) \quad \text{for } \delta = 1-2mc_1\tilde\gamma^2 \min\{1,\tfrac{\gamma^2}{2M}\}<1 .$$

↖ linear convergence ↖ depends on the condition number $\frac{m}{M}$ of $D^2 f$!



proof: • $f(x_k+\alpha \underset{p_k^S}{\underset{\|}{p_k}}) \le f(x_k)+\alpha Df(x_k)p_k + \frac{M}{2}\alpha^2 \underbrace{\|p_k\|_2^2}_{\le \|p_k\|^2/\gamma^2 = \|\nabla f(x_k)\|_*^2/\gamma^2 = -Df(x_k)p_k/\gamma^2} = f(x_k) + \alpha(1-\frac{M}{2\gamma^2}\alpha) Df(x_k)p_k$

$\Rightarrow$ any $\alpha\in(0,\frac{2\gamma^2}{M}(1-c_1)) \supset (0,\gamma^2/M)$ satisfies Armijo's condition $\Rightarrow \alpha_k \ge \min\{1,\frac{\gamma^2}{2M}\}$

• $f(x_{k+1}) = f(x_k+\alpha_k p_k) \le f(x_k) + c_1\alpha_k Df(x_k)p_k \le f(x_k) - c_1\min\{1,\frac{\gamma^2}{2M}\}\|\nabla f(x_k)\|_*^2$

$\le f(x_k) - c_1\tilde\gamma^2\min\{1,\frac{\gamma^2}{2M}\} \underbrace{\|\nabla f(x_k)\|_2^2}_{\ge 2m(f(x_k)-p^*)}$

$\Rightarrow f(x_{k+1})-p^* \le f(x_k)-p^* - 2mc_1\tilde\gamma^2\min\{1,\frac{\gamma^2}{2M}\}(f(x_k)-p^*) = \delta(f(x_k)-p^*)$ $\square$

· What does this imply about $\|Df(x_k)\|_2$, $\|x_k-x^*\|_2$?

· What rate follows from Zoutendijk's Thm?

Acceptance of Newton step

why important?

Thm: If $\cdot\, D^2 f$ continuous $\cdot\, mI \le D^2 f(x) \le MI \;\forall x$ $\cdot\, c_1 < \frac{1}{2}$ $\cdot \lim_{k\to\infty} \|p_k\| = 0$ $\cdot \lim_{k\to\infty} \frac{\|p_k^N - p_k\|}{\|p_k\|} = 0$

then for k sufficiently large, $\alpha_k = 1$ satisfies the strong Wolfe conditions.

proof: 0) $$\frac{Df(x_k)p_k}{\|p_k\|_2^2} + m \le \frac{Df(x_k)p_k}{\|p_k\|_2^2} + \frac{p_k^T D^2 f(x_k) p_k}{\|p_k\|_2^2} = \frac{p_k^T D^2 f(x_k)(p_k - p_k^N)}{\|p_k\|_2^2} \le M \frac{\|p_k - p_k^N\|_2}{\|p_k\|_2} \to 0$$

$\Rightarrow$ for k large enough, $-Df(x_k)p_k \ge \frac{m}{2}\|p_k\|_2^2$

1) Armijo's condition satisfied for k large enough:

$\cdot$ Taylor's thm: $f(x_k + p_k) = f(x_k) + Df(x_k)p_k + \frac{1}{2} p_k^T D^2 f(x_k + q_k) p_k$ for some $q_k \in [0, p_k]$

$\cdot$ $$f(x_k + p_k) - f(x_k) - \frac{1}{2} Df(x_k)p_k = \frac{1}{2}\left[Df(x_k)p_k + p_k^T D^2 f(x_k + q_k) p_k\right]$$

$$= \frac{1}{2}\Big[\underbrace{\big(Df(x_k)p_k + p_k^T D^2 f(x_k) p_k^N\big)}_{=0} + \underbrace{p_k^T D^2 f(x_k)(p_k - p_k^N)}_{\le M\, o(\|p_k\|_2^2)} + \underbrace{p_k^T\big(D^2 f(x_k + q_k) - D^2 f(x_k)\big)p_k}_{= o(\|p_k\|_2^2)}\Big] = o\left[\left(c_1 - \frac{1}{2}\right) Df(x_k)p_k\right]$$

2) strong Wolfe condition satisfied for k large enough:

$$|Df(x_k + p_k)p_k| = \Big|\underbrace{Df(x_k)p_k + p_k^T D^2 f(x_k) p_k}_{= p_k^T D^2 f(x_k)(p_k - p_k^N) \le M\, o(\|p_k\|_2^2)} + \underbrace{\int_0^1 p_k^T\big(D^2 f(x_k + t p_k) - D^2 f(x_k)\big)p_k \, dt}_{= o(\|p_k\|_2^2)}\Big|,$$

while $-c_2 Df(x_k)p_k \ge c_2 \frac{m}{2}\|p_k\|_2^2$ $\square$

Convergence rate Newton's method

Thm: Let $D^2 f$ Lipschitz with constant L, $mI \leq D^2 f(x) \leq MI$, $c_1 < \frac{1}{2}$,

$p_k = p_k^N$, $\alpha_k = 1$ where possible. Then $\limsup_{k\to\infty} \frac{\|x_{k+1} - x^*\|_2}{\|x_k - x^*\|_2^2} \leq \frac{L}{m}$. quadratic convergence

proof: • Zoutendijk $\Rightarrow$ $\|Df(x_k)\|_2 \to 0$ $\Rightarrow$ $\|x_k - x^*\|_2 \to 0$ & $p_k \to 0$

• By previous theorem, $\alpha_k = 1$ for k large enough, i.e. $x_{k+1} = x_k + p_k$

• $D^2 f(x_k)(x_{k+1} - x^*) = D^2 f(x_k)(p_k - (x^* - x_k)) = Df(x^*) - Df(x_k) - D^2 f(x_k)(x^* - x_k)$

$$= \int_0^1 \left(D^2 f(x_k + t(x^* - x_k)) - D^2 f(x_k)\right)(x^* - x_k)\, dt$$

• upon taking norms on both sides,

$m\|x_{k+1} - x^*\|_2 \leq \|D^2 f(x_k)(x_{k+1} - x^*)\|_2 \leq \left\|\int_0^1 (D^2 f(x_k + t(x^* - x_k)) - D^2 f(x_k))(x^* - x_k)\, dt\right\|_2 \leq L\|x_k - x^*\|_2^2$ □

Rem: • If $mI \leq D^2 f(x^*) \leq MI$, then by continuity, analogous bounds hold in a neighbourhood.

• If the algorithm at some point reaches such a neighbourhood, we have quadratic convergence

• If $D^2 f$ is only Hölder-continuous with exponent $\alpha > 0$?

• Newton's method is invariant under coordinate transforms $x \mapsto Ax$ (check!), but analysis is not $\Rightarrow$ actually, only $\|(D^2 f(x)^{-1} D^2 f(x + t(x^* - x)) - Id)(x^* - x)\|_2 \leq C\|x^* - x\|_2^2$ required, i.e. no separate Lipschitz & strong convexity condition (related to "self-concordance")

Complexity of Newton's method

<u>Thm:</u> Under same conditions as before, $\underbrace{\|Df(x_k)\|_2 < \min\left(\frac{1}{2}-c_1, \frac{c_2}{2}\right)\frac{m^2}{L}}_{\text{implies } \|x_k - x^*\|_2 < \min\left(\frac{1}{2}-c_1, c_2\right)\frac{m}{L}}$ implies $\|x_{k+1}-x^*\|_2 \leq \frac{L}{m}\|x_k - x^*\|_2^2$.

proof: same proof holds, only need to show $\alpha_k = 1$; for this repeat previous proof, using

$$p_k^{N^T}\left(D^2 f(x_k+q_k) - D^2 f(x_k)\right)p_k^N \leq L\|p_k^N\|_2^3 \leq L\frac{\|Df(x_k)\|_2}{m}\|p_k^N\|_2^2 \leq \begin{cases} \left(\frac{1}{2}-c_1\right)m\|p_k^N\|_2^2 \leq \left(c_1 - \frac{1}{2}\right)Df(x_k)p_k^N \\ \frac{c_2}{2}m\|p_k^N\|_2^2 \end{cases}$$

$\square$

<u>Thm:</u> Under same conditions as before and assuming α_k to be chosen maximally up to a factor $\beta \in (0,1)$, for all $\eta > 0$ there is a $\gamma > 0$ s.t. $\|Df(x_k)\|_2 \geq \eta$ implies $f(x_{k+1}) \leq f(x_k) - \gamma$.

proof: $\cdot$ $-Df(x_k)p_k^N = p_k^{N^T} D^2 f(x_k) p_k^N \geq m\|p_k^N\|_2^2$

$\cdot$ $f(x_k + \alpha p_k^N) \leq f(x_k) + \alpha Df(x_k)p_k^N + \frac{M}{2}\alpha^2\|p_k^N\|_2^2 \leq f(x_k) + \alpha\left(1 - \frac{\alpha}{2}\frac{M}{m}\right)Df(x_k)p_k^N$

$\Rightarrow \hat{\alpha} = 2\frac{m}{M}(1-c_1)$ satisfies Armijo's condition

$\cdot$ case 1 ($\hat{\alpha}$ satisfies Wolfe conditions): $\alpha_k \geq \beta\hat{\alpha} \Rightarrow f(x_k + \alpha_k p_k^N) - f(x_k) \leq \beta\hat{\alpha}c_1 \underbrace{Df(x_k)p_k^N}_{\leq -\frac{1}{M}\eta^2} \leq -\gamma$

case 2 ($Df(x_k + \hat{\alpha}p_k^N)p_k^N$ too negative): $\alpha_k \geq \hat{\alpha} \geq \beta\hat{\alpha}$, —— " —— $\gamma \stackrel{!}{=} \frac{\beta\hat{\alpha}c_1\eta^2}{M}$

case 3 ($Df(x_k + \hat{\alpha}p_k^N)p_k^N$ too positive): $f(x_k + \alpha_k p_k^N) - f(x_k) \leq f(x_k + \hat{\alpha}p_k^N) - f(x_k) \leq -\gamma$. $\square$

<u>Cor:</u> Newton's method reaches $\|x_k - x^*\|_2 < \varepsilon$ after $\frac{f(x_0) - p^*}{\gamma} + \log_2\left(\frac{\log \varepsilon}{\log \min\left(\frac{1}{2} - c_1, \frac{c_2}{2}\right)}\right)$ iterations.

Modified Newton's method

- If $D^2 f(x_k)$ is indefinite, one can choose $p_k = -(D^2 f(x_k) + M_k)^{-1} \nabla f(x_k)$ for M_k such that $D^2 f(x_k) + M_k$ is position definite, e.g.

$$M_k = \max(0, \varepsilon - \lambda_{min}) I \text{ for } \lambda_{min} \text{ the smallest eigenvalue of } D^2 f(x_k).$$

- In quasi-Newton methods we approximate the Hessian just from the past gradients.

Thm: Same conditions as before, only $\lim_{k\to\infty} \frac{\|p_k - p_k^N\|_2}{\|p_k\|_2} = 0$. Then $\limsup_{k\to\infty} \frac{\|x_{k+1} - x^*\|_2}{\|x_k - x^*\|_2} = 0$.

superlinear convergence

proof: • As before, $\|x_k - x^*\|_2 \to 0$ & $p_k \to 0$ & $x_{k+1} = x_k + p_k$

• $$\underbrace{D^2 f(x_k)(x_{k+1} - x^*)}_{\text{norm} \geq m \|x_{k+1} - x^*\|_2} = D^2 f(x_k)(x_k - x^* + p_k^N - p_k^N + p_k) = \underbrace{Df(x^*) - Df(x_k) - D^2 f(x_k)(x^* - x_k)}_{\text{norm} \leq L \|x_k - x^*\|_2^2} + \underbrace{D^2 f(x_k)(p_k - p_k^N)}_{\text{norm} \leq M\, o(\|p_k\|_2)}$$

$$\Rightarrow m \|x_{k+1} - x^*\|_2 \leq L \|x_k - x^*\|_2^2 + M\, o(\underbrace{\|p_k\|_2}_{= \|x_{k+1} - x_k\|_2 \leq \|x_{k+1} - x^*\|_2 + \|x_k - x^*\|_2})$$

$$\Rightarrow \|x_{k+1} - x^*\|_2 \leq o(\|x_k - x^*\|_2) \qquad \square$$

Quasi-Newton methods

• Abbreviate: $g_k = \nabla f(x_k)$, $s_k = x_{k+1} - x_k$, $y_k = g_{k+1} - g_k$, H_k = approximation of $D^2 f(x_k)$

Rem: Taylor $\Rightarrow \nabla f(x_k) = \nabla f(x_{k+1}) + D^2 f(x_{k+1})(x_k - x_{k+1})$ + h.o.t. $\Leftrightarrow D^2 f(x_{k+1}) s_k = y_k$ + h.o.t.

Def: The condition $H_{k+1} s_k = y_k$ for $H_{k+1} \in \mathbb{R}^{n \times n}$ is called the secant condition.

secant with slope H_{k+1} "secant method": $x_{k+1} = x_k - \alpha_k \left(\frac{f'(x_k) - f'(x_{k-1})}{x_k - x_{k-1}} \right)^{-1} f'(x_k)$

Rem: If H_{k+1} is positive definite, the secant condition requires $s_k^T y_k > 0$.

Thm: A descent step satisfying the strong Wolfe conditions satisfies $s_k^T y_k > 0$.

proof: strong Wolfe $\Rightarrow g_{k+1}^T s_k \geq c_2 g_k^T s_k \Rightarrow y_k^T s_k \geq \underbrace{(c_2 - 1)}_{<0} \alpha_k \underbrace{g_k^T p_k}_{<0} > 0$ $\square$

Quasi-Newton methods choose $p_k = -H_k^{-1} g_k$, where H_0 is an initial approximation of $D^2 f(x_0)$ (typically $H_0 = I$), and H_k is an improved approximation, which is updated in each step to satisfy the secant condition. Typically, the update is of low rank; different updates lead to different methods.

Low-rank updates

Thm (Sherman-Morrison-Woodbury formula): Let $A \in \mathbb{R}^{n\times n}$, $U, V \in \mathbb{R}^{n\times p}$.

$$B = A + UV^T \Rightarrow B^{-1} = A^{-1} - A^{-1}U\left(I + V^TA^{-1}U\right)^{-1}V^TA^{-1}.$$

proof: By calculation (homework) $\square$

Ex: $\left(I + \begin{pmatrix}1\\2\\3\end{pmatrix}(1\ 2\ 3)\right)^{-1} v = \begin{pmatrix}2 & 2 & 3\\ 2 & 5 & 6\\ 3 & 6 & 10\end{pmatrix}^{-1} v = Iv - \begin{pmatrix}1\\2\\3\end{pmatrix}\left(1 + (1\ 2\ 3)\begin{pmatrix}1\\2\\3\end{pmatrix}\right)^{-1}(1\ 2\ 3)v = v - \frac{\begin{pmatrix}1\\2\\3\end{pmatrix}\cdot v}{15}\begin{pmatrix}1\\2\\3\end{pmatrix}$

Def: • The symmetric rank-1 (SR1) update is given by

$$H_{k+1} = H_k + \frac{(y_k - H_k s_k)(y_k - H_k s_k)^T}{(y_k - H_k s_k)^T s_k}$$ rank-1 update

• The Davidon-Fletcher-Powell (DFP) update is given by

$$H_{k+1} = \left(I - \frac{y_k s_k^T}{y_k^T s_k}\right) H_k \left(I - \frac{s_k y_k^T}{y_k^T s_k}\right) + \frac{y_k y_k^T}{y_k^T s_k}$$ rank-2 update

• The Broyden-Fletcher-Goldfarb-Shanno (BFGS) update is given by

$$H_{k+1} = H_k - \frac{H_k s_k s_k^T H_k}{s_k^T H_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}$$ rank-2 update

Rem: By the S-M-W formula,

$$B_{k+1} = H_{k+1}^{-1} = \begin{cases} B_k + \dfrac{(s_k - B_k y_k)(s_k - B_k y_k)^T}{(s_k - B_k y_k)^T y_k} & \text{(SR 1)} \\ B_k - \dfrac{B_k y_k y_k^T B_k}{y_k^T B_k y_k} + \dfrac{s_k s_k^T}{y_k^T s_k} & \text{(DFP)} \\ \left(I - \dfrac{s_k y_k^T}{y_k^T s_k}\right) B_k \left(I - \dfrac{y_k s_k^T}{y_k^T s_k}\right) + \dfrac{s_k s_k^T}{y_k^T s_k} & \text{(BFGS)} \end{cases}$$

SR1, Davidon-Fletcher-Powell, Broyden-Fletcher-Goldfarb-Shanno

Rem: • SR1 is the unique symmetric rank-1 update satisfying the secant condition (homework)

- The SR1 update can fail. (zero denominator)
- Even if H_k is pos. def., H_{k+1} need not be. (homework)

• DFP is the rank-2 update satisfying the secant condition and being "closest" to H_k:

DFP solves $H_{k+1} = \arg\min_H \|G^{-\frac{1}{2}}(H-H_k)G^{-\frac{1}{2}}\|_F^2$ s.t. $H=H^T$, $Hs_k = y_k$; $G = \int_0^1 D^2 f(x_k + \tau(x_{k+1}-x_k))\,d\tau$

- Under the strong Wolfe conditions, the DFP update is well-defined, since $y_k^T s_k > 0$.
- H_k pos. def. $\Rightarrow$ H_{k+1} pos. def.

(indeed, for $z \in \mathbb{R}^n$ let $w = \left(I - \frac{s_k y_k^T}{y_k^T s_k}\right) z$, then $z^T H_{k+1} z = w^T H_k w + \frac{(y_k^T z)^2}{y_k^T s_k} \geq 0$,

and $z^T H_{k+1} z = 0 \Rightarrow y_k^T z = 0 \Rightarrow w = z$ & $z^T H_{k+1} z = z^T H_k z > 0$ ↯)

• BFGS has the same properties as DFP, only for B_{k+1} instead of H_{k+1}.

Convergence of BFGS

note: estimate of $\cos\theta_k$ via eigenvalues of Hessian is here replaced by an estimate via $\operatorname{trace}(H_k)$, $\det(H_k)$

Thm: If $f \in C^2(\mathbb{R}^n)$ with $mI \le D^2 f \le MI$, then $x_k \to x^*$ for the iterates x_k of the BFGS quasi-Newton method with $H_0 = I$ and strong Wolfe stepsize control.

proof:

- Define $m_k = \frac{y_k^T s_k}{s_k^T s_k}$, $M_k = \frac{\|y_k\|_2^2}{y_k^T s_k}$, $q_k = \frac{s_k^T H_k s_k}{s_k^T s_k}$, $\cos\theta_k = \frac{s_k^T H_k s_k}{\|s_k\|_2 \|H_k s_k\|_2} = \frac{-g_k^T p_k}{\|g_k\|_2 \|p_k\|_2}$
- note: $m_k \ge m$, $M_k \le M$
- $\operatorname{trace}(H_{k+1}) = \operatorname{trace}(H_k) - \frac{\|H_k s_k\|_2^2}{s_k^T H_k s_k} + \frac{\|y_k\|_2^2}{y_k^T s_k} = \operatorname{trace}(H_k) - \frac{q_k}{\cos^2\theta_k} + M_k$ (homework)
- $\det(H_{k+1}) = \det(H_k) \frac{y_k^T s_k}{s_k^T H_k s_k} = \det(H_k)\frac{m_k}{q_k}$ (homework)
- Let $\psi(H) = \operatorname{trace} H - \ln(\det H)$, then $\psi(H) > 0$ for H pos. def. (homework)

and $\psi(H_{k+1}) = \psi(H_k) + \underbrace{(M_k - \ln m_k - 1)}_{\le M - \ln m - 1 = c \ge 0} + \underbrace{\left[1 - \frac{q_k}{(\cos\theta_k)^2} + \ln\frac{q_k}{\cos^2\theta_k}\right]}_{\le 0} + \ln\cos^2\theta_k$

$\Rightarrow 0 < \psi(H_{k+1}) \le \psi(H_0) + c(k+1) + \sum_{j=0}^{k} \ln\cos^2\theta_j$

- assume $\cos\theta_k \to 0$, then $\exists k_1 > 0\ \forall k > k_1: \ln\cos^2\theta_k \le -2c$ (wlog, $c > 0$)

$\Rightarrow 0 < \psi(H_0) + c(k+1) + \sum_{j=0}^{k_1} \ln\cos^2\theta_j - 2c(k-k_1) \xrightarrow[k\to\infty]{} -\infty$ ↯

Hence, $\cos\theta_k$ does not converge to $0 \underset{\text{Zoutendijk}}{\Rightarrow} Df(x_k) \to 0$ for subsequence $\underset{\text{strong convexity}}{\Rightarrow} x_k \to x^*$. □

Linear convergence of BFGS

Thm: Under the same conditions, $\exists \mu \in (0,1), \rho > 0$: $\|x_k - x^*\|_2 \leq \rho \mu^k \|x_0 - x^*\|_2$.

proof: • already know $0 < \psi(H_0) + c(k+1) + \sum_{j=0}^{k} \ln \cos^2 \theta_j$

• $\forall r \in (0,1) \ \exists \varkappa > 0$: $\cos^2 \theta_j \geq \varkappa$ for at least $\lfloor r(k+1) \rfloor$ indices j in $\{0, \dots, k\}$ $\forall k$.

Indeed, let $\eta_j = -\ln \cos^2 \theta_j$ and let η^k be the $\lfloor r(k+1) \rfloor$ th smallest of $\{\eta_0, \dots, \eta_k\}$.

Set $I^k = \{ i \in \{0, \dots, k\} \mid \eta_i > \eta^k \}$. For $i \in \{0, \dots, k\} \setminus I^k$,

$$\eta_i \leq \eta^k \leq \frac{1}{|I^k|} \sum_{j \in I^k} \eta_j \leq \frac{1}{1-r} \frac{1}{k+1} \sum_{j=0}^{k} \eta_j < \frac{1}{1-r} \left[\frac{\psi(H_0)}{k+1} + c \right] \leq C$$

$\Rightarrow \cos \theta_i \geq \exp(-C/2) =: \sqrt{\varkappa}$

• recall from proof of Zoutendijk's theorem: $f(x_k) - f(x_{k+1}) \geq c_1 \frac{1-c_2}{M} \cos^2 \theta_k \underbrace{\|Df(x_k)\|_2^2}_{\geq 2m (f(x_k) - f(x^*))}$

$\Rightarrow f(x_{k+1}) - f(x^*) \leq \left(1 - c_1 \frac{1-c_2}{M} 2m \cos^2 \theta_k\right) \left(f(x_k) - f(x^*)\right)$

$\Rightarrow f(x_{k+1}) - f(x^*) \leq \left(1 - c_1 \frac{1-c_2}{M} 2m \varkappa\right)^{\lfloor r(k+1) \rfloor} \left(f(x_0) - f(x^*)\right)$

monotonicity of $f(x_k)$ and $\cos^2 \theta_k \geq \varkappa$ for fraction r of all indices

• result follows from $\frac{1}{M} \left(f(x) - f(x^*)\right) \leq \|x - x^*\|_2^2 \leq \frac{1}{m} \left(f(x) - f(x^*)\right)$

for $\rho = \left(\frac{M}{m}\right)^{1/2}$, $\mu = \left(1 - c_1 \frac{1-c_2}{M} 2m \varkappa\right)^{r/2}$ $\square$

Superlinear convergence of BFGS

for simplicity, we rescale domain s.t. $D^2 f(x^*) = I$

Thm: Under same conditions as before, if $D^2 f$ is Lipschitz with constant L, $\lim_{k\to\infty} \frac{\|x_{k+1}-x^*\|_2}{\|x_k - x^*\|_2} = 0$.

proof: • Let $G = D^2 f(x^*)$, $\tilde s_k = G^{\frac12} s_k$, $\tilde y_k = G^{-\frac12} y_k$, $\tilde H_k = G^{-\frac12} H_k G^{-\frac12}$

• as before, define $\tilde m_k = \frac{\tilde y_k^T \tilde s_k}{\tilde s_k^T \tilde s_k}$, $\tilde M_k = \frac{\|\tilde y_k\|_2^2}{\tilde y_k^T \tilde s_k}$, $\tilde q_k = \frac{\tilde s_k^T \tilde H_k \tilde s_k}{\tilde s_k^T \tilde s_k}$, $\cos\tilde\theta_k = \frac{\tilde s_k^T \tilde H_k \tilde s_k}{\|\tilde s_k\|_2 \|\tilde H_k \tilde s_k\|_2}$

• as before, $\tilde H_{k+1} = \tilde H_k - \frac{\tilde H_k \tilde s_k \tilde s_k^T \tilde H_k}{\tilde s_k^T \tilde H_k \tilde s_k} + \frac{\tilde y_k \tilde y_k^T}{\tilde y_k^T \tilde s_k}$, $\Psi(\tilde H_{k+1}) = \Psi(\tilde H_k) + (\tilde M_k - \ln \tilde m_k - 1) + \left[1 - \frac{\tilde q_k}{\cos^2\tilde\theta_k} + \ln\frac{\tilde q_k}{\cos^2\tilde\theta_k}\right] + \ln\cos^2\tilde\theta_k$

• $y_k - G s_k = \int_0^1 D^2 f(x_k + \tau s_k) - G \, d\tau \, s_k$

$\Rightarrow \|\tilde y_k - \tilde s_k\|_2 \le \underbrace{\|G^{-\frac12}\|_2}_{\le 1/\sqrt{m}} \|y_k - G s_k\|_2 \le \sqrt{\frac{1}{m}} L \varepsilon_k \|s_k\|_2 \le \frac{1}{m} L \varepsilon_k \|\tilde s_k\|_2$ for $\varepsilon_k = \max(\|x_k - x^*\|_2, \|x_{k+1}-x^*\|_2)$

• we have $\tilde y_k^T \tilde s_k = (\tilde s_k + (\tilde y_k - \tilde s_k))^T \tilde s_k \ge \|\tilde s_k\|_2^2 \left(1 - \frac{L}{m}\varepsilon_k\right)$, $\|\tilde y_k\|_2^2 \le \|\tilde s_k + (\tilde y_k - \tilde s_k)\|_2^2 \le \|\tilde s_k\|_2^2 \left(1 + \frac{L}{m}\varepsilon_k\right)^2$

$\Rightarrow \ln \tilde m_k \ge \ln\left(1 - \frac{L}{m}\varepsilon_k\right) \ge -c\varepsilon_k$; $\tilde M_k \le \frac{\left(1+\frac{L}{m}\varepsilon_k\right)^2}{1-\frac{L}{m}\varepsilon_k} \le 1 + c\varepsilon_k$ for some $c > 0$

$\Rightarrow 0 < \Psi(\tilde H_{k+1}) \le \Psi(\tilde H_k) + 3c\varepsilon_k + \ln\cos^2\tilde\theta_k + \left[1 - \frac{\tilde q_k}{\cos^2\tilde\theta_k} + \ln\frac{\tilde q_k}{\cos^2\tilde\theta_k}\right]$ use previous thm & geometric series

$\Rightarrow \sum_{j=0}^\infty \left(\ln\frac{1}{\cos^2\tilde\theta_j} - \left[1 - \frac{\tilde q_j}{\cos^2\tilde\theta_j} + \ln\frac{\tilde q_j}{\cos^2\tilde\theta_j}\right]\right) \le \Psi(\tilde H_0) + 3c\sum_{j=0}^\infty \varepsilon_j < \infty$, i.e. $\lim_{j\to\infty}\cos\tilde\theta_j = 1$, $\lim_{j\to\infty}\tilde q_j = 1$

• $\frac{\|G^{-\frac12}(H_k - G)s_k\|_2^2}{\|G^{1/2}s_k\|_2^2} = \frac{\|(\tilde H_k - I)\tilde s_k\|_2^2}{\|\tilde s_k\|_2^2} = \frac{\tilde q_k^2}{\cos^2\tilde\theta_k} - 2\tilde q_k + 1 \xrightarrow{k\to\infty} 0$ now apply earlier convergence result!

$\Rightarrow \frac{\|p_k - p_k^N\|_2}{\|p_k\|_2} \le \frac{\|D^2 f(x_k)(p_k - p_k^N)\|_2}{m\|p_k\|_2} = \frac{\|(D^2 f(x_k) - H_k)p_k\|_2}{m\|p_k\|_2} \le \frac{\|(H_k - G)p_k\|_2 + \|(G - D^2 f(x_k))p_k\|_2}{m\|p_k\|_2} \xrightarrow{k\to\infty} 0$ □

Simple treatment of linear constraints

<u>general</u>: $\min_{x\in\mathbb{R}^n} f(x)$ s.t. $Ax=b$; $f\in C^2(\mathbb{R}^n)$ convex, $A\in\mathbb{R}^{p\times n}$, rank $A=p<n$ (P)

- KKT condition: $\exists\lambda\in\mathbb{R}^p$ s.t. $\begin{cases} Ax^*=b \\ \nabla f(x^*)+A^T\lambda=0 \end{cases}$

example on next slide!

Elimination of constraints: • find $F\in\mathbb{R}^{n\times(n-p)}$ with range $F=\ker A$ & $\hat{x}$ with $A\hat{x}=b$

• (P) $\Longleftrightarrow$ $\min_{z\in\mathbb{R}^{n-p}} f(Fz+\hat{x})$ (unconstrained optim.)

<u>quadratic</u>: $\min_{x\in\mathbb{R}^n} \frac{1}{2}x^TPx+q^Tx+r$ s.t. $Ax=b$; $P\in\mathbb{R}^{n\times n}$ sym. pos. semi-def., $q\in\mathbb{R}^n$, $r\in\mathbb{R}$

- KKT condition: $\exists\lambda\in\mathbb{R}^p$ s.t. $\underbrace{\begin{pmatrix} P & A^T \\ A & 0 \end{pmatrix}}_{\text{KKT matrix}}\begin{pmatrix} x^* \\ \lambda \end{pmatrix}=\begin{pmatrix} -q \\ b \end{pmatrix}$

  - if KKT matrix regular $\Rightarrow$ $\exists!$ primal-dual optimal pair (x^*,λ)
  - if KKT matrix singular & $\exists$ solution $\Rightarrow$ every solution is primal dual optimal pair
  - if KKT matrix singular & there is no solution $\Rightarrow$ $p^*=-\infty$

Projection methods for linear constraints

$\min_{x\in\mathbb{R}^n} f_0(x)$ s.t. $Ax=b$, $A\in\mathbb{R}^{p\times n}$, $\operatorname{rank} A = p < n$

Thm: The orthogonal projection of $d\in\mathbb{R}^n$ onto the tangent space $T=\{s\in\mathbb{R}^n \mid As=0\}$ to the set of feasible points is given by $\operatorname{proj}_T d = (I - A^T(AA^T)^{-1}A)d$.

proof: write $d = d_1 + d_2$, $d_1 = \operatorname{proj}_T d \in \ker A$, $d_2 \in (\ker A)^\perp = \operatorname{range} A^T$

$\Rightarrow d = d_1 + A^T s$ for some $s\in\mathbb{R}^p$ $\Rightarrow Ad = (AA^T)s \Rightarrow s = (AA^T)^{-1}Ad \Rightarrow d_1 = d - A^T(AA^T)^{-1}Ad$ $\square$

Rem: $-\operatorname{proj}_T \nabla f_0(x)$ is a descent direction, since $(-\operatorname{proj}_T \nabla f_0(x))\cdot\nabla f_0(x) = -\|\operatorname{proj}_T \nabla f_0(x)\|_2^2$.

Def: The linesearch methods with descent direction

$$\left.\begin{aligned}
p_k &= \operatorname*{argmin}_{p\in\mathbb{R}^n,\ Ap=0,\ p\cdot p\le Df(x_k)\cdot p} f_0(x_k) + Df(x_k)p \\
p_k^N &= \operatorname*{argmin}_{p\in\mathbb{R}^n,\ Ap=0} f_0(x_k) + Df(x_k)p + \frac{1}{2}p^T D^2 f(x_k)p
\end{aligned}\right\}$$

minimise first-/second order Taylor approximation under constraints

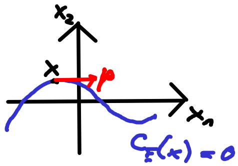
are called <u>projected gradient/Newton descent</u>, respectively. The directions are computed via

$$p_k = -\operatorname{proj}_T \nabla f_0(x_k) \quad , \quad \begin{pmatrix} D^2 f_0(x_k) & A^T \\ A & 0 \end{pmatrix}\begin{pmatrix} p_k^N \\ w \end{pmatrix} = \begin{pmatrix} -\nabla f_0(x_k) \\ 0 \end{pmatrix}.$$

Projection methods for nonlinear constraints

$$\min_{x \in \mathbb{R}^n} f_0(x) \quad \text{s.t.} \quad c_E(x) = 0$$

<u>problem</u>: At a feasible point x there is in general no feasible direction p (a direction s.t. $c_E(x + \alpha p) = 0$ for α small).



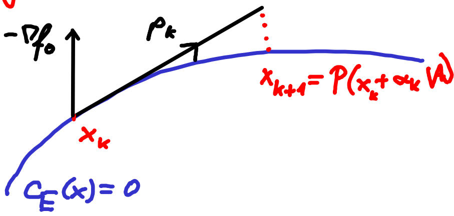
<u>Alg. (projection method):</u>

given: $f_0 : \mathbb{R}^n \to \mathbb{R}$, $x_0 \in \mathbb{R}^n$, $c_E(x_0) = 0$

repeat

1) choose a descent direction $p_k \in T = \{ p \in \mathbb{R}^n \mid Dc_E(x_k) p = 0 \}$ — tangent space to feasible set; e.g. projected gradient/Newton step
2) choose a step length $\alpha_k > 0$
3) $x_{k+1} = P(x_k + \alpha_k p_k)$ — projection onto $\{ c_E(x) = 0 \}$
4) $k \leftarrow k + 1$

until x_{k+1} sufficiently minimises f_0



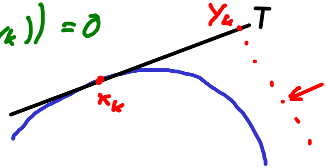
<u>Rem</u>:
- For the projection P there are several possibilities.
- To choose a stepsize α_k one can use the same criteria as earlier, applied to $f_0 \circ P$.

Projection onto constraint set

From a point $y_k = x_k + \alpha_k p_k$ close to x_k, need to find a point $P(y_k) \in \{c_E(x) = 0\}$

- choose $P(y_k)$ on the constraint set in a direction orthogonal to the tangent plane T at x_k

$\Rightarrow P(y_k) = y_k + \nabla c_E(x_k)\, s$ for $s \in \mathbb{R}^p$ s.t. $c_E(P(y_k)) = 0$



here, $P(y_k)$ not defined!

- always possible for y_k close to x_k

- perform Newton's method to find a zero s of $c_E(\underbrace{y_k + \nabla c_E(x_k)s}_{=:y_s}) =: h(s)$, i.e.

$$s_0 = 0,\quad s_{i+1} = s_i - Dh(s_i)^{-1} h(s_i) = s_i - \left(Dc_E(y_{s_i})\, Dc_E(x_k)^T\right)^{-1} c_E(y_{s_i}), \text{ i.e.}$$

$$y_{s_0} = y_k,\quad y_{s_{i+1}} = y_{s_i} - Dc_E(x_k)^T \left[Dc_E(y_{s_i})\, Dc_E(x_k)^T\right]^{-1} c_E(y_{s_i})$$

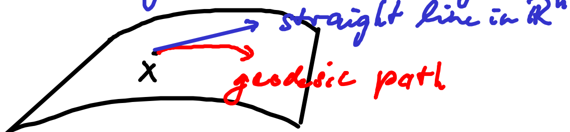
- alternatively, use $Dc_E(y_{s_i}) \approx Dc_E(x_k)$ $\Rightarrow$ $y_{s_{i+1}} = y_{s_i} - Dc_E(x_k)\left[Dc_E(x_k)\, Dc_E(x_k)^T\right]^{-1} c_E(y_{s_i})$

same matrices as for gradient projection

<u>Rem:</u> The reduced gradient method introduced later can be seen as a projection method using a different projection P.

Convergence rate of proj. grad. desc.: preliminaries

Idea:
- analyze simplified algorithm that asymptotically duplicates the considered method
- let $S = \{x \in \mathbb{R}^n \mid c_E(x) = 0\}$, $T_x S$ = tangent plane to S in x
- imagine a bug living on S minimising f_0 via gradient descent; at each iteration
  - the bug chooses the (projected) gradient descent direction p_k and
  - moves in that direction along a straight line with step length α_k
- "straight lines" in S are so-called geodesics; if iterates x_k converge against x^*, then the difference between geodesics and straight lines in $\mathbb{R}^n$ converges to 0 even faster



$\Rightarrow$ the bug's iterates are asymptotically the same as the projected gradient descent ones

$\Rightarrow$ the bug's descent has the same convergence rate

Convergence rate of proj. grad. desc.: geodesics & Lagrangians

Def: • The arclength of a smooth curve $x:[0,T]\to S$ is given by $l[x]=\int_0^T |\dot x(t)|\,dt$.

• A geodesic is a curve $x(t)$ minimising $l[x]$ for fixed $x(0), x(T)$. (Ex: great circles on sphere)

We shall assume $|\dot x(t)|=1 \ \forall t\in[0,T]$ (which can always be achieved).

Rem: Geodesics satisfy $Dc_E(x(t))\dot x(t)=0$ (from $0=\frac{d}{dt}c_E(x(t))$)

and $\ddot x(t)=\nabla c_E(x(t))\,\omega(t)$ for some $\omega:[0,T]\to\mathbb{R}^p$ (optimality condition for $\min_x l[x]$)

acceleration only normal, not tangential to S

The above two differential equations together with $|\dot x(0)|=1$ can be shown to uniquely define a curve $x(t)$, a geodesic, with $|\dot x(t)|=1$ as long as $x(t)$ is regular.

Thm: Let $L_x(x,\lambda)=Df_0(x)+\lambda^T D\underbrace{c_E}_{=(g_1(x),\dots,g_p(x))^T}(x)$ and $L_{xx}(x,\lambda)=D^2f_0(x)+\sum_{i=1}^p\lambda_i D^2g_i(x)$ denote derivative and Hessian of the Lagrangian and introduce the Lagrange multiplier $\lambda(x)=-[Dc_E(x)Dc_E(x)^T]^{-1}Dc_E(x)\nabla f_0(x)$

Along a geodesic $x(t)$ on S, $\frac{d}{dt}f_0(x(t))=L_x(x,\lambda(x))\dot x$, $\frac{d^2}{dt^2}f_0(x(t))=\dot x^T L_{xx}(x,\lambda(x))\dot x$.

proof: • $\frac{d}{dt}f_0(x(t))=Df_0(x)\dot x(t)=\mathrm{proj}_{T_xS}\nabla f_0(x)\cdot\dot x(t)=(I-Dc_E^T[Dc_E\,Dc_E^T]^{-1}Dc_E)\nabla f_0\cdot\dot x=L_x(x(t),\lambda(x(t)))\dot x$

• $\frac{d^2}{dt^2}f_0(x(t))=\dot x^T D^2f_0(x)\dot x+Df_0(x)\ddot x$; also $\tilde\lambda^T c_E(x(t))=0 \Rightarrow 0=\dot x(t)^T\sum_{i=1}^p\tilde\lambda_i D^2g_i(x(t))\dot x(t)+\tilde\lambda^T Dc_E(x(t))\ddot x(t)$

adding both equations, $\frac{d^2}{dt^2}f_0(x(t))=\dot x^T L_{xx}(x,\tilde\lambda)\dot x+\underbrace{(Df_0(x)+\tilde\lambda^T Dc_E(x))\ddot x}$ $\perp\ddot x$ for $\tilde\lambda=\lambda(x)$ □

Optimization algorithms: Projection methods for equality-constrained optim.

Convergence rate of proj. grad. desc. (assume $f \in C^2$)

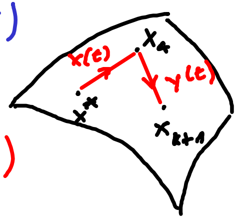
<u>Thm</u>: Let the iterates x_k of geodesic gradient descent converge to x^* (where for simplicity we take α_k as the minimising step length), and let $mI \le L_{xx}(x,\lambda(x)) \le MI$ for x close to x^*, then asymptotically, $\frac{f(x_{k+1})-f(x^*)}{f(x_k)-f(x^*)} \le \left(\frac{M-m}{M+m}\right)^2$.

proof: abbreviate $g(t) = L_x(x(t),\lambda(x(t)))^T$, $H(t) = L_{xx}(x(t),\lambda(x(t)))$, $g_k = L_x(x_k,\lambda(x_k))^T$, $H_k = L_{xx}(x_k,\lambda(x_k))$

- let $x(t)$ be a geodesic with $x(0)=x^*$, $x(\underset{}{\overset{= \text{dist}(x^*,x_k)}{T}}) = x_k$, then by Taylor expansion of $f_0(x(t))$ & $g(t)$,

$$f_0(x^*) - f_0(x_k) = -g_k \dot x(T) T + \frac{1}{2} T^2 \dot x^T(T) H_k \dot x(T) + o(T^2) \qquad (*)$$

$$g_k = g(T) - g(0) = \dot g(T) T + o(T) = T H_k \dot x(T) + T \nabla c_E(x_k) \frac{d}{dt}(\lambda \circ x)(T) + o(T)$$



- let P_k the orthogonal projection onto $T_{x_k}S$, then $g_k = P_k g_k = P_k H_k \dot x(T) T + o(T)$

$\xRightarrow{\dot x_k = P_k \dot x_k}$ $\tilde H_k \dot x(T) T = g_k + o(T)$ for $\tilde H_k = P_k H_k P_k$ $\left(\xRightarrow{|\dot x_k|=1} mT \le |g_k| + o(T) \le MT\right)$

$\xRightarrow{\text{in }(*)}$ $f_0(x^*) - f_0(x_k) = -\frac{1}{2} g_k^T \tilde H_k^{-1} g_k + o(T^2) = -\frac{1}{2} g_k^T \tilde H_k^{-1} g_k (1 + o(1))$

- now let $y(t)$ be the geodesic with $y(0) = x_k$, $\dot y(0) = \frac{p_k}{\|p_k\|_2} = -\frac{g_k}{\|g_k\|_2}$, then

$$f_0(y(t)) = f_0(x_k) + t g_k^T \dot y(0) + \frac{t^2}{2} \dot y^T(0) H_k \dot y(0) + o(t^2)$$ is minimised by $t_k = \frac{-g_k^T \dot y}{\dot y^T H_k \dot y} + o(t_k) \sim T$

$$\Rightarrow f_0(x_k) - f_0(x_{k+1}) = \frac{1}{2}\frac{(g_k^T \dot y)^2}{\dot y H_k \dot y} + o(T^2) = \frac{1}{2}\frac{(g_k^T \dot y)^2}{\dot y H_k \dot y}(1 + o(1)) = \frac{1}{2}\frac{\|g_k\|_2^4}{g_k^T \tilde H_k g_k}(1 + o(1))$$

Kantorovich inequality

- $f_0(x_{k+1}) - f_0(x^*) = f_0(x_k) - f_0(x^*) + f_0(x_{k+1}) - f_0(x_k) = [f_0(x_k) - f_0(x^*)]\left(1 - \frac{\|g_k\|_2^4 (1 + O(1))}{g_k^T \tilde H_k g_k \; g_k^T \tilde H_k^{-1} g_k}\right) \le [f_0(x_k) - f_0(x^*)]\left(1 - \frac{4 + O(1)}{(\sqrt{m/M} + \sqrt{M/m})^2}\right)$ $\square$

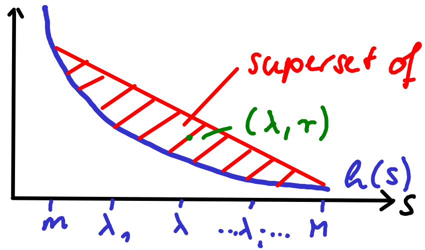
The Kantorovich inequality

Thm: Let $p_1,\dots,p_n \geq 0$, $0 < m \leq \lambda_1,\dots,\lambda_n \leq M$, then $\left(\sum_{i=1}^n p_i\lambda_i\right)\left(\sum_{i=1}^n p_i/\lambda_i\right) \leq \frac{(m+M)^2}{4mM}\left(\sum_{i=1}^n p_i\right)^2$

proof: $\xi_1 = p_1/\sum_{i=1}^n p_i, \dots, \xi_n = p_n/\sum_{i=1}^n p_i$ are convex combination coefficients (i.e. $\xi_i \in [0,1]$, $\sum_{i=1}^n \xi_i = 1$)

$h(s) = 1/s$

$\Rightarrow (\lambda, r) := \sum_{i=1}^n \xi_i (\lambda_i, h(\lambda_i))$ lies in the convex hull of the $(\lambda_i, h(\lambda_i))$



$$\Rightarrow r \leq \frac{\lambda - m}{M-m} h(M) + \frac{M-\lambda}{M-m} h(m) = \frac{m+M-\lambda}{mM}$$

$$\lambda_{max} = \frac{m+M}{2}$$

$$\Rightarrow \left(\sum_{i=1}^n p_i\lambda_i\right)\left(\sum_{i=1}^n \frac{p_i}{\lambda_i}\right)\Big/\left(\sum_{i=1}^n p_i\right)^2 = \left(\sum_{i=1}^n \xi_i\lambda_i\right)\left(\sum_{i=1}^n \frac{\xi_i}{\lambda_i}\right) = \lambda r \leq \lambda\frac{m+M-\lambda}{mM} \leq \max_{\lambda\in\mathbb{R}} \lambda\frac{m+M-\lambda}{mM} \overset{!}{=} \frac{(m+M)^2}{4mM} \qquad \square$$

Cor: Let $0 < mI \leq H \leq MI$ for a symmetric positive definite $H \in \mathbb{R}^{n\times n}$, $x \in \mathbb{R}^n$, then $\frac{(x^T H x)(x^T H^{-1} x)}{\|x\|_2^4} \leq \frac{(m+M)^2}{4mM}$.

proof: Let $\lambda_1,\dots,\lambda_n \in [m,M]$ be the eigenvalues of H with orthonormal eigenvectors $q_1,\dots,q_n \in \mathbb{R}^n$.

Write $x = \sum_{i=1}^n z_i q_i$, then $\|x\|_2^2 = \sum_{i=1}^n z_i^2$, $(x^T H x)(x^T H^{-1} x) = \left(\sum_{i=1}^n \lambda_i z_i^2\right)\left(\sum_{i=1}^n \frac{1}{\lambda_i} z_i^2\right)$, set $p_i = z_i^2$ $\square$

Rem: Note $\frac{(m+M)^2}{4mM} = \frac{1}{4}\left(\sqrt{\frac{m}{M}} + \sqrt{\underbrace{\frac{M}{m}}_{\text{condition number}}}\right)^2 = \frac{1}{1 - \left(\frac{M-m}{M+m}\right)^2}$.

Linear constraints

$$\min_{x\in\mathbb{R}^n} f_0(x) \quad \text{s.t. } Ax=b \qquad , \quad A\in\mathbb{R}^{p\times n}, \ \operatorname{rank} A = p < n$$

choose a basis B of $\mathbb{R}^p$ from the columns of A (for simplicity let B contain the first p columns) and subdivide $\quad x=(y,z)\in\mathbb{R}^p\times\mathbb{R}^{n-p}$, $\quad A=[B,C]\in\mathbb{R}^{p\times p}\times\mathbb{R}^{p\times(n-p)}$

Rem: z can be regarded as the independent and $y=B^{-1}(b-Cz)$ as the dependent variable.

Def: $\tilde f(z) = f_0(\underbrace{B^{-1}(b-Cz)}_{y(z)}, z)$ is called the reduced functional,
$\nabla\tilde f(z) = \left[D_z f_0(y(z),z) - D_y f_0(y(z),z)B^{-1}C\right]^T$ the reduced gradient,
and $D^2\tilde f(z)$ the reduced Hessian.

The reduced gradient/Newton method is a gradient/Newton descent for $\tilde f$.

Nonlinear constraints

$$\min_{x\in\mathbb{R}^n} f_0(x) \quad \text{s.t.} \quad c_E(x)=0$$

- Let $x=(y,z)\in\mathbb{R}^p\times\mathbb{R}^{n-p}$ and assume $\nabla_y c_E(y,z)$ to be regular $\xRightarrow{\text{implicit fct thm}}$ $y=y(z)$ locally
- Define reduced functional $\tilde{f}(z)=f_0(y(z),z)$.

$$\left.\begin{aligned} &c_E(y(z),z)=0 \Rightarrow D_y c_E(y(z),z)\frac{dy}{dz}+D_z c_E(y(z),z)=0 \\ &\nabla\tilde{f}(z) = \left[D_y f_0(y(z),z)\frac{dy}{dz}+D_z f_0(y(z),z)\right]^T \end{aligned}\right\} \Rightarrow \begin{aligned} \nabla\tilde{f}(z) = {}&\nabla_z f_0(y(z),z) - \\ &\nabla_z c_E(y(z),z)\nabla_y c_E(y(z),z)^{-1}\nabla_y f_0(y(z),z)\end{aligned}$$

- analogously for reduced Hessian
- The reduced gradient/Newton method requires to find $y(z_{k+1})$ for any given z_{k+1}.

As for projection methods, this can be done by a Newton iteration to find a zero of $h(y)=c_E(y,z_{k+1})$,

$$y_0=y(z_k), \qquad y_{i+1}=y_i - D_y c_E(y_i,z_{k+1})^{-1} c_E(y_i,z_{k+1})$$

or alternatively, using $D_y c_E(y_i,z_{k+1})\approx D_y c_E(y(z_k),z_k)$

$$y_{i+1}=y_i - D_y c_E(y(z_k),z_k)^{-1} c_E(y_i,z_{k+1})$$

Convergence rate of reduced gradient descent

<u>Thm</u>: Abbreviate $K(z) = \frac{dy}{dz} = -D_y c_E(y(z),z)^{-1} D_z c_E(y(z),z)$ and $C(z) = \begin{pmatrix} K(z) \\ I \end{pmatrix} \in \mathbb{R}^{n \times (n-p)}$, then for any $\lambda \in \mathbb{R}^r$ satisfying $D_y f_0(y(z),z) + \lambda^T D_y c_E(y(z),z) = 0$ we have $D^2 \tilde{f}(z) = C(z)^T L_{xx}((y(z),z),\lambda)\, C(z)$.

proof: $\cdot\ \tilde{f}(z) = f_0(y(z),z) = f_0(y(z),z) + \lambda^T c_E(y(z),z)$

$\cdot\ D\tilde{f}(z) = [D_y f_0(y(z),z) + \lambda^T D_y c_E(y(z),z)] K(z) + D_z f_0(y(z),z) + \lambda^T D_z c_E(y(z),z) = \overbrace{L_y((y(z),z),\lambda)}^{=0} K(z) + L_z((y(z),z),\lambda)$

$\cdot\ D^2\tilde{f}(z) = K(z)^T L_{yy}((y(z),z),\lambda) K(z) + L_{zy}((y(z),z),\lambda) K(z) + K(z)^T L_{yz}((y(z),z),\lambda) + L_{zz}((y(z),z),\lambda)$ □

<u>Thm</u>: Let the iterates x_k of reduced gradient descent converge to x^* (where for simplicity we take α_k as the minimising step length), and let $mI \leq C(z)^T L_{xx}(x,\lambda(x)) C(z) \leq MI$ for x close to x^*, then asymptotically, $\frac{f(x_{k+1}) - f(x^*)}{f(x_k) - f(x^*)} \leq \left(\frac{M-m}{M+m}\right)^2$.

proof: homework □

<u>Rem</u>: The convergence rate of the projected gradient descent depends on the eigenvalues of $\tilde{H} = P^T L_{xx}(x^*, \lambda(x^*)) P$ for P the orthogonal projection onto $T_{x^*}S$. We can write $P = C(z^*)(C(z^*)^T C(z^*))^{-\frac{1}{2}}$ (indeed, the columns of C span $T_{x^*}S$ since $(D_y c_E(x^*)\ D_z c_E(x^*))\, C = 0$, thus the columns of P form an orthonormal basis of $T_{x^*}S$ due to $P^T P = I$). Hence, the convergence rate of the reduced gradient descent depends on the eigenvalues of $\sqrt{C^T C}\, \tilde{H} \sqrt{C^T C}$ and thus is coordinate-dependent depends on chosen z

Idea of interior point methods

min $f_0(x)$ s.t. $f_1(x), \dots, f_m(x) \le 0$, $Ax = b$, $f_i \in C^2$, convex, $A \in \mathbb{R}^{p \times n}$, rank $A = p < n$ (P)

<u>assume:</u> · optimal point x^* exists

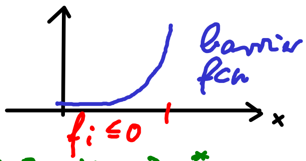
· Slater's condition (strict feasibility) satisfied

$\Rightarrow$ strong duality, KKT conditions fulfilled

$$\left.\begin{array}{l} Ax^* = b, \ f_i(x^*) \le 0, \ \mu_i^* \ge 0, \ \mu_i^* f_i(x^*) = 0, \ i = 1, \dots, m \\ \nabla f_0(x^*) + \sum_{i=1}^{m} \mu_i^* \nabla f_i(x^*) + A^T \lambda^* = 0 \end{array}\right\} \quad (O)$$

<u>idea:</u> · Solve (P) or (O) by applying Newton's method to a sequence of modifications of (P) or (O) that only have equality constraints.

· E.g., replace $f_i(x^*) \le 0$ by adding a "barrier function" to f_0



· interior point methods produce a sequence x_k with $Ax_k = b$, $f_i(x_k) < 0$, $x_i \to x^*$

Barrier methods: barrier functions

(P) $\Leftrightarrow$ $\min_x f_0(x) + \sum_{i=1}^m \chi_-(f_i(x))$ s.t. $Ax = b$ for $\chi_-(s) = \begin{cases} 0 & \text{if } s \le 0 \\ \infty & \text{else} \end{cases}$

replace $\chi_-(u)$ by $I_t(u) = -\frac{1}{t}\ln(-u)$, $\operatorname{dom} I_t = \{u \in \mathbb{R} \mid u < 0\}$

<u>Thm</u>: For $t > 0$ • I_t is convex, differentiable, non-decreasing

• $I_t(u) \xrightarrow[u \uparrow 0]{} \infty$, $I_t(u) = \infty$ for $u > 0$ $\square$

<u>Def</u>: • The <u>logarithmic barrier function</u> ϕ to (P) is given by

$$\phi(x) = -\sum_{i=1}^m \ln(-f_i(x)), \quad \operatorname{dom}\phi = \{x \in \mathbb{R}^n \mid f_1(x), \dots, f_m(x) < 0\}$$

$$\nabla\phi(x) = -\sum_{i=1}^m \frac{1}{f_i(x)} \nabla f_i(x), \quad D^2\phi(x) = \sum_{i=1}^m \frac{1}{f_i^2(x)} \nabla f_i(x) \nabla f_i(x)^T - \frac{1}{f_i(x)} D^2 f_i(x)$$

ϕ is convex.

• The <u>barrier problem</u> to (P) is given by

$$\min_x \; t f_0(x) + \phi(x) \quad \text{s.t. } Ax = b \qquad (P_t)$$

• The <u>central path</u> is $\{x^*(t) \mid t > 0\}$, where $x^*(t)$ is the optimal point of (P_t); it satisfies $\exists \lambda \in \mathbb{R}^p$: $0 = t\nabla f_0 + \nabla\phi + A^T\lambda$

Barrier methods: algorithm

Thm: Let (x^*, λ^*) be optimal primal dual pair for (P_t), then (μ^*, λ^*) is dually feasible for $\mu_i^* = \frac{-1}{t f_i(x^*)}$, and $f_0(x^*) - g(\mu^*, \lambda^*) = \frac{m}{t}$, i.e. x^* and (μ^*, λ^*) are $\frac{m}{t}$-suboptimal.

proof: • $\mu_i^* > 0$ due to $f_i(x^*) < 0$

• optimality condition for (P_t): $0 = \nabla f_0 + \frac{\nabla \phi}{t} + A^T \lambda^* = \nabla f_0(x^*) + \sum_{i=1}^m \mu_i^* \nabla f_i(x^*) + A^T \lambda^*$

$\Rightarrow x^* = \operatorname{argmin}_x \left(f_0(x) + \sum_{i=1}^m \mu_i^* f_i(x) + \lambda^{*T}(Ax - b) \right) = \operatorname{argmin}_x L(x, (\mu^*, \lambda^*))$

• $p^* \geq g(\mu^*, \lambda^*) = \min_x L(x, (\mu^*, \lambda^*)) = L(x^*, (\mu^*, \lambda^*)) = f_0(x^*) - \frac{1}{t} \sum_{i=1}^m \frac{f_i(x^*)}{f_i(x^*)}$ □

Alg (barrier method): given: strictly feasible x, $t = t_0 > 0$, $\beta > 1$, tolerance $\varepsilon > 0$

iterate until $\frac{m}{t} < \varepsilon$:

step 1: solve (P_t) via Newton's method with initial guess x

step 2: $x \leftarrow x^*(t)$, $t \leftarrow \beta t$

Rem: The previous theorem implies convergence $f_0(x^*(t)) \to p^*$.

Barrier methods: complexity

usually not fulfilled for barrier functions but can be relaxed to a "self-concordance" assumption

Thm: Let $D^2 f_0 + \frac{1}{t} D^2 \phi$ Lipschitz with constant L, $\tilde{m} I \le D^2 f_0(x) + \frac{1}{t} D^2 \phi \le M I$ for some $\tilde{m}, M > 0$ and all $t > 0$, then the barrier method reaches an x with $\|x - x^*\|_2 < \varepsilon$ after $N = \frac{\log \frac{m}{t_0 \varepsilon}}{\log \beta} \left(\frac{m(\beta - 1 - \log \beta)}{\gamma} + C \log_2 \log \frac{1}{\tilde{\varepsilon}} \right)$ steps for some $C, \gamma > 0$, where $\tilde{\varepsilon}$ is the accuracy of solving (P_t).

linear convergence

proof:
- every iteration takes $\frac{f(x) - p^*}{\gamma} + C \log_2 \log \frac{1}{\tilde{\varepsilon}}$ Newton steps starting from x where $f = t f_0 + \phi$, $x = x^*(t/\beta)$, $x^+ = x^*(t)$, $p^* = f(x^+)$
- $t f_0(x) + \phi(x) - t f_0(x^+) - \phi(x^+) = t f_0(x) - t f_0(x^+) + \sum_{i=1}^m \left(-\log(-f_i(x)) + \log(-f_i(x^+)) \right)$

$$= t f_0(x) - t f_0(x^+) + \sum_{i=1}^m \underbrace{\log(-t \mu_i f_i(x^+))}_{\le -t\mu_i f_i(x^+) - 1} - m \log \beta$$

$\mu_i := \frac{\beta}{-t f_i(x)}$ (multiplier from previous iteration)

$$\le t f_0(x) - t \underbrace{\left(f_0(x^+) + \sum_{i=1}^m \mu_i f_i(x^+) + \lambda^T \underbrace{(A x^+ - b)}_{0} \right)}_{\ge g(\mu, \lambda) = f_0(x) - \frac{m}{t}\beta} - m - m \log \beta \le m(\beta - 1 - \log \beta)$$

(similar to duality gap on previous slide!)

$\Rightarrow$ need less than $\frac{m(\beta - 1 - \log \beta)}{\gamma} + C \log_2 \log \frac{1}{\tilde{\varepsilon}}$ Newton steps

- in total $\log \frac{m}{t_0 \varepsilon} / \log \beta$ outer iterations

$\square$

Barrier methods: finding a feasible starting point

Barrier methods require a strictly feasible starting point x_0 with $c_I(x_0) = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}(x_0) < 0$, $Ax_0 = b$.

This is found by an auxiliary optimisation, a so-called phase I method. the actual optimisation problem is then called phase II

method 1: $\min\limits_{x \in \mathbb{R}^n, s \in \mathbb{R}} \; s \quad$ s.t. $\; f_1(x), \dots, f_m(x) \le s \;, \; Ax = b$

- strictly feasible optimisation problem, just choose some x with $Ax = b$ and $s > \max\limits_i f_i(x)$

  $\Rightarrow$ apply barrier method

- $p^* < 0 \Rightarrow$ feasible x_0 exists (stop phase I method as soon as $s < 0$ and take x as x_0)

  $p^* > 0 \Rightarrow$ original problem infeasible

  $p^* = 0 \Rightarrow$ original problem $\begin{cases} \text{not feasible if minimum not attained} \\ \text{not strictly feasible else} \Rightarrow \text{barrier method not applicable} \end{cases}$

method 2: $\min\limits_{x \in \mathbb{R}^n, s \in \mathbb{R}^m} \; s_1 + \dots + s_m \quad$ s.t. $\; f_1(x) \le s_1, \dots, f_m(x) \le s_m \;, \; Ax = b \;, \; s_1, \dots, s_m \ge 0$

- also strictly feasible $\Rightarrow$ apply barrier method

- if $p^* > 0$, at least some $f_i(x)$ are > 0; only those constraints $f_i(x) \le 0$ for which $s_i > 0$ (and corresponding dual variables are 0) are violated

Barrier methods: phase I termination near phase II central path

Assume we know that a feasible point can be found with $f_0(x) \le M$, then choose

phase I : $\min_{x \in \mathbb{R}^n, s \in \mathbb{R}} \quad s \qquad \text{s.t.} \quad f_1(x), \dots, f_m(x) \le s, \quad f_0(x) \le M, \quad Ax = b$

- The central path $(s(t), x(t))$ for the logarithmic barrier method applied to this satisfies the optimality conditions $\quad t = \sum_{i=1}^m \frac{1}{s - f_i(x)}, \quad 0 = \frac{1}{M - f_0(x)} \nabla f_0(x) + \sum_{i=1}^m \frac{1}{s - f_i(x)} \nabla f_i(x) + A^T \lambda$

- The central path $x(t)$ for the barrier method applied to the original problem satisfies
$$t \nabla f_0(x) + \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla f_i(x) + A^T \lambda = 0$$

- $\Rightarrow$ If (s, x) lies on the central path of phase I with $s = 0$, then x lies on the central path of phase II with parameter $t = \frac{1}{M - f_0(x)}$ and duality gap $m(M - f_0(x))$

Barrier methods: complexity of phase I

- Assume, we know the feasible set to lie in a ball of radius R (for simplicity no eq. constraints)
- Consider phase I problem $\min\limits_{x\in\mathbb{R}^n,\, s\in\mathbb{R}} s$ s.t. $f_1(x),\dots,f_m(x)\le s$, $a^Tx\le 1$

  where a will be chosen with $\|a\|_2\le\frac{1}{R}$ ($\Rightarrow a^Tx\le 1$ is redundant); start with $s=s_0$, $x=0$
- Choose a & s_0 such that $(s=s_0, x=0)$ lies on central path for phase I, i.e.

$$t_0=\sum_{i=1}^m\frac{1}{s_0-f_i(0)} \quad\text{and}\quad a=-\sum_{i=1}^m\frac{1}{s_0-f_i(0)}\nabla f_i(0).$$

  Also need $\|a\|_2\le\frac{1}{R}$: $\|a\|_2\le\sum_{i=1}^m\frac{1}{s_0-f_i(0)}\|\nabla f_i(0)\|_2\le\frac{mG}{s_0-F}$ for $G=\max\limits_i\|\nabla f_i(0)\|_2$, $F=\max\limits_i f_i(0)$

  $\Rightarrow$ choose $s_0=mGR+F$

- Initial duality gap $\frac{m+1}{t_0}=(m+1)mGR\Big/\sum_{i=1}^m\frac{1}{1+(F-f_i(0))/(mGR)}\le(m+1)mGR$,

  note: phase I has m+1 constraints

  needed accuracy $\varepsilon=|p^*|$ (= measure of difficulty of determining feasibility)

  only need to determine whether $p^*>0$ or $p^*<0$

  $\Rightarrow$ number of Newton steps needed:

  take $\beta=1+\frac{1}{\sqrt{m+1}}$, $\log(1+x)=x-\frac{1}{2}\left(\frac{x}{z}\right)^2$ for some $z\in[1,1+x]$

$$N=\underbrace{\frac{\log\frac{m+1}{t_0\varepsilon}}{\log\beta}\left(\frac{(m+1)(\beta-1-\log\beta)}{\gamma}+C\log_2\log\frac{1}{\varepsilon}\right)}_{\text{tends to }\infty\text{ for }\beta\to1\text{ or }\beta\to\infty}\le C_1\sqrt{m+1}\,\log\left(\frac{m(m+1)GR}{|p^*|}\right)\left(\frac{1}{2\gamma}+c\right)$$

$\Rightarrow$ there is optimal β

Primal dual interior point methods

Similar to barrier methods, but :
- no inner/outer iterations
- search directions from Newton's method applied to KKT conditions
- primal & dual iterates not necessarily feasible
- can exhibit superlinear convergence
- can be applied if problem not strictly feasible

Rem: In the logarithmic barrier method, the Newton step p_k^N of the inner iterations solves

$$\begin{pmatrix} D^2 f_0(x) + \frac{1}{t} D^2\phi(x) & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} p_k^N \\ \lambda \end{pmatrix} = - \begin{pmatrix} \nabla f_0(x) + \frac{1}{t}\nabla\phi(x) \\ 0 \end{pmatrix}. \qquad (*)$$

This can be interpreted as a Newton step for solving the modified KKT system

$$\begin{aligned} Df_0(x) + \sum_{i=1}^m \mu_i Df_i(x) + \lambda^T A &= 0 \\ -\mu_i f_i(x) &= \frac{1}{t}, \quad i = 1,\dots,m \\ Ax - b &= 0 \end{aligned} \qquad (mKKT)$$

Indeed, eliminating $\mu_i = -1/(t f_i(x))$, the above becomes $Df_0(x) + \frac{1}{t} D\phi(x) + \lambda^T A = 0$, $Ax - b = 0$

for which $(*)$ is a Newton step starting from an (x, λ) with $Ax = b$, $\lambda = 0$.

The primal-dual search direction will be a Newton step for (mKKT) in (x, μ, λ).

Primal dual interior point methods: algorithm

<u>Def</u>: · The <u>primal dual step</u> is a Newton step for (mKKT), i.e., $p_k^{pd} = (\Delta x, \Delta\mu, \Delta\lambda)$ with

$$\begin{pmatrix} D^2 f_0(x) + \sum_{i=1}^m \mu_i D^2 f_i(x) & \nabla c_I(x) & A^T \\ -\mathrm{diag}(\mu_1,\dots,\mu_m) Dc_I(x) & -\mathrm{diag}(f_1(x),\dots,f_m(x)) & 0 \\ A & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\mu \\ \Delta\lambda \end{pmatrix} = -\begin{pmatrix} r_{dual} \\ r_{central} \\ r_{primal} \end{pmatrix} := -\begin{pmatrix} \nabla f_0(x) + \nabla c_I(x)\mu + A^T\lambda \\ -\mathrm{diag}(\mu_1,\dots,\mu_m) c_I(x) - \frac{1}{t} \\ Ax - b \end{pmatrix}$$

for the <u>dual, central, and primal residual</u> r_{dual}, $r_{central}$, r_{primal}.

· The <u>surrogate duality gap</u> is $\eta(x,\mu) = -\mu^T c_I(x)$.

<u>Rem</u>: Since x need not be feasible (unlike in barrier methods), we cannot compute a duality gap, but

$$\left.\begin{array}{ll} r_{dual}=0,\ \mu\geq 0 & \Rightarrow x = \operatorname{argmin} L(x,(\mu,\lambda)) \\ r_{primal}=0 & \Rightarrow Ax=b \end{array}\right\} \Rightarrow p^* - d^* \leq f_0(x) - L(x,(\mu,\lambda)) = \eta(x,\mu)$$

then surrogate duality gap is true duality gap

<u>Alg</u>: given x_0 with $f_1(x_0),\dots,f_m(x_0) < 0$, $\mu_0 > 0$, λ_0, $\beta > 1$, ε_{feas}, $\varepsilon > 0$

repeat 1) $t := \frac{\beta m}{\eta(x_k,\mu_k)}$ 2) compute p_k^{pd} 3) choose step size $\alpha_k > 0$ 4) $\begin{pmatrix} x_{k+1} \\ \mu_{k+1} \\ \lambda_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ \mu_k \\ \lambda_k \end{pmatrix} + \alpha_k p_k^{pd}$

until $\|r_{primal}\|_2 < \varepsilon_{feas}$, $\|r_{dual}\|_2 < \varepsilon_{feas}$, $\eta(x_k,\mu_k) < \varepsilon$

<u>Rem</u>: If x_k, μ_k, λ_k solved (mKKT), η would be the duality gap and $\frac{m}{\eta}$ the barrier parameter t $\Rightarrow$ step 1) means $t \leftarrow \beta t$

Primal dual interior point methods: linesearch

Step 3) can be done via backtracking, using Armijo's condition, i.e., starting from $\alpha_k = 1$,

repeat $\alpha_k \leftarrow \frac{\alpha_k}{2}$ until
$$\begin{cases} c_I(x_k + \alpha_k \Delta x) < 0 \\ \mu_k + \alpha_k \Delta\mu_k > 0 \\ \| r((x_k, \mu_k, \lambda_k) + \alpha_k p_k^{pd}) \|_2 \le (1 - c_1 \alpha_k) \| r(x_k, \mu_k, \lambda_k) \|_2 \end{cases}$$

$(r_{primal}, r_{central}, r_{dual})$

Thm: The backtracking terminates.

proof: • $c_I < 0$ and $\mu > 0$ are satisfied for α_k small enough by continuity.

• Let $y = (x_k, \mu_k, \lambda_k)$, then $p_k^{pd} = -Dr(y)^{-1} r(y)$

$$\Rightarrow \frac{d}{dt} \| r(y + t p_k^{pd}) \|_2^2 \Big|_{t=0} = 2 r(y)^T Dr(y) p_k^{pd} = -2 r(y)^T r(y)$$

$$\Rightarrow \frac{d}{dt} \sqrt{\| r(y + t p_k^{pd}) \|_2^2} \Big|_{t=0} = -\| r(y) \|_2 \quad \Rightarrow \| r(y + t p_k^{pd}) \|_2 = (1 - t) \| r(y) \|_2 + O(t^2) \quad \square$$

Thm: Once a full Newton step is taken, $r_{primal} = 0$ (i.e., $Ax_k = b$) for all following iterations.

proof: In that step, $Ax_k = b$; but then, all following directions Δx are feasible. $\square$

Thm: In each step, r_{primal} decreases by the factor $(1 - \alpha_k)$.

proof: $Ax_{k+1} - b = Ax_k - b + \alpha_k A \Delta x = (1 - \alpha_k)(Ax_k - b)$ $\square$

Penalty methods

$$\min_{x \in \mathbb{R}^n} f_0(x) \quad \text{s.t.} \quad c_I(x) = (f_1(x), \dots, f_m(x)) \le 0 \qquad (P)$$

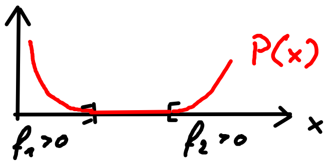
replace potential equality constraints $g_i(x) = 0$ by $g_i(x) \le 0$ & $-g_i(x) \le 0$

<u>Def</u>: A <u>penalty function</u> for the constraints $f_1(x), \dots, f_m(x) \le 0$ is a function $P(x) = \gamma(c_I^+(x))$ with $c_I^+(x) = (\max\{0, f_1(x)\}, \dots, \max\{0, f_m(x)\})$ and $\gamma : [0,\infty)^m \to [0,\infty)$ such that

$$\gamma(0) = 0, \quad \gamma(p) > 0 \text{ for } p \neq 0, \quad \gamma \text{ continuous}$$

<u>Ex</u>: · $\gamma(p) = \frac{1}{2} p^T \Gamma p$ for a positive definite $\Gamma \in \mathbb{R}^{m \times m}$

· case $\Gamma = I$ yields $P(x) = \frac{1}{2} \sum_{i=1}^m (\max\{0, f_i(x)\})^2$



<u>Alg</u> (penalty method): given $0 < c_1 < c_2 < \dots$, $\lim_{k \to \infty} c_k = \infty$

for $k = 1, 2, \dots$ do $\quad x_k = \arg\min_x q(c_k, x) = \arg\min_x f_0(x) + c_k P(x)$

We shall assume a solution x_k to exist for each k.

Penalty methods: convergence

Thm: We have $q(c_k, x_k) \le q(c_{k+1}, x_{k+1})$, $P(x_k) \ge P(x_{k+1})$, $f_0(x_k) \le f_0(x_{k+1})$.

proof: $q(c_{k+1}, x_{k+1}) = f_0(x_{k+1}) + c_{k+1} P(x_{k+1}) \ge f_0(x_{k+1}) + c_k P(x_{k+1}) \ge f_0(x_k) + c_k P(x_k) = q(c_k, x_k)$

- $\boxed{f_0(x_k) + c_k P(x_k) \le f_0(x_{k+1}) + c_k P(x_{k+1})} + \boxed{f_0(x_{k+1}) + c_{k+1} P(x_{k+1}) \le f_0(x_k) + c_{k+1} P(x_k)} = \boxed{(c_{k+1} - c_k) P(x_{k+1}) \le (c_{k+1} - c_k) P(x_k)}$
- $f_0(x_{k+1}) + c_k P(x_{k+1}) \ge f_0(x_k) + c_k P(x_k) \ge f_0(x_k) + c_k P(x_{k+1}) \Rightarrow f_0(x_{k+1}) \ge f_0(x_k)$ □

Thm: Let x^* solve (P), then $f_0(x^*) \ge q(c_k, x_k) \ge f_0(x_k)$.

proof: $f_0(x^*) = f_0(x^*) + c_k P(x^*) \ge f_0(x_k) + c_k P(x_k) \ge f_0(x_k)$ □

Thm: Let $x_1, x_2, \ldots$ be generated by the penalty method, then any limit point solves (P).

proof:
- Let $K \subset \mathbb{N}$ indicate a subsequence with $\lim_{k \in K} x_k = x$ (thus $\lim_{k \in K} f_0(x_k) = f_0(x)$)
- $q(c_k, x_k)$ is monotone & bounded by $f_0(x^*)$, thus $\lim_{k \in K} q(c_k, x_k) = q^* \le f_0(x^*)$
- $\lim_{k \in K} c_k P(x_k) = \lim_{k \in K} q(c_k, x_k) - f_0(x_k) = q^* - f_0(x) \Rightarrow P(x) = \lim_{k \in K} P(x_k) = 0$ x feasible!
- $f_0(x) = \lim_{k \in K} f_0(x_k) \le f_0(x^*)$ by above thm x optimal! □

Active set methods

$$\min_{x \in \mathbb{R}^n} f_0(x) \quad \text{s.t.} \quad f_1(x), \dots, f_m(x) \le 0, \quad g_1(x) = \dots = g_p(x) = 0 \qquad (P)$$

<u>Def</u>: Given the optimal point $x^* \in \mathbb{R}^n$, the <u>active set</u> is defined as $A = \{ i \in \{1, \dots, m\} \mid f_i(x^*) = 0 \}$.

If A were known, (P) would reduce to an equality-constrained optimisation.
Active set methods guess and update A (the current guess is called "working set" W)

<u>Alg (active set method)</u>: Given $f_0, \dots, f_m, g_1, \dots, g_p$, $W \subset \{1, \dots, m\}$

repeat

1) minimise $f_0(x)$ s.t. $f_i(x) = g_j(x) = 0 \quad \forall j \in \{1, \dots, p\}, i \in W \qquad (P_W)$

   if $\underbrace{f_i(x) > 0}_{\text{constraint was violated}}$ for some $i \notin W$, add i to W and repeat step 1)

2) update working set: $W = \{ j \in \{1, \dots, m\} \mid \overbrace{f_j = 0 \text{ and } \mu_j \ge 0}^{\text{constraint is active}} \}$

   Lagrange multiplier for constraint $f_j = 0$

until W does not change

Active set methods: convergence

Rem: · Step 1) can be performed via any method for equality-constrained optimisation.

· When W does not change, then for some $\mu \in \mathbb{R}^m$, $\lambda \in \mathbb{R}^p$ we have

$$0 = Df_0(x) + \sum_{i=1}^{m} \mu_i Df_i(x) + \sum_{i=1}^{p} \lambda_i D g_i(x) \quad , \quad g_1(x) = \ldots = g_p(x) = 0 \, ,$$

$$f_i(x) < 0 \, , \mu_i = 0 \, , \, i \notin W \quad , \quad f_i(x) = 0 \, , \mu_i \geq 0 \, , \, i \in W ,$$

i.e. $W = A$ and x satisfies the KKT conditions.

Thm: Suppose, for every $W \subset \{1, \ldots, m\}$ the problem (P_W) has a unique nondegenerate solution (i.e. $\mu_i \neq 0 \; \forall i \in W$). Then the active set method converges to the optimal point of (P)

proof: After the solution for one working set W is found, a decrease in the objective is made

$\Rightarrow$ it is impossible to return to that working set.

Process must terminate with $W = A$ since there are only finitely many working sets. $\square$

Rem: · Many solves with incorrect W needed.

· To update W, in principle, (P_W) needs to be solved exactly. could have infinitely many changes of W

· In practice, one updates (P_W) prematurely (based on heuristics), but then convergence unclear.

Local convexity

$f_0, \dots, f_m, g_1, \dots, g_p \in C^2$

$$\min_{x \in \mathbb{R}^n} f_0(x) \quad \text{s.t.} \quad c_E(x) = (g_1, \dots, g_p)^T(x) = 0, \quad c_I(x) = (f_1, \dots, f_m)^T(x) \le 0 \qquad (P)$$

Let x^* be a local minimum of (P) & let it be a regular point wrt. the constraints.

- $\exists\, \lambda^* \in \mathbb{R}^p, \mu^* \in [0,\infty)^m$: $\quad Df_0(x^*) + \lambda^{*T} Dc_E(x^*) + \mu^{*T} Dc_I(x^*) = 0$, $\quad \mu_i^* f_i(x^*) = 0$, and
- $L_{xx}(x^*, \mu^*, \lambda^*) = D^2 f_0(x^*) + \sum_{i=1}^p \lambda_i^* D^2 g_i(x^*) + \sum_{i=1}^m \mu_i^* D^2 f_i(x^*)$ is positive semi-definite on the tangent space $T_{x^*}S$ to the active constraints (i.e. $c_E(x) = 0$ and $f_i(x) \le 0$ for $i \in J = \{ i \in \{1, \dots, m\} \mid f_i(x^*) = 0 \}$)

<u>Def</u>: We call $\quad L_{xx}(x^*, \mu^*, \lambda^*)$ is positive definite on $\mathbb{R}^n$ $\qquad$ (C)

the <u>local convexity assumption.</u>

<u>Thm</u>: Under (C), x^* is a local minimum of the unconstrained problem $\min_{x \in \mathbb{R}^n} f_0(x) + \lambda^{*T} c_E(x) + \mu^{*T} c_I(x)$.

Also, for any (μ, λ) in a neighbourhood of (μ^*, λ^*) with $\mu_i = 0$ for $i \notin J$, $\underbrace{f_0(x) + \lambda^T c_E(x) + \mu^T c_I(x)}_{L(x,\mu,\lambda)}$ has a local minimum x near x^*.

proof: $\cdot$ x^* satisfies the second order sufficient conditions for optimality

- $L_{xx}(x^*, \mu^*, \lambda^*)$ pos. def. $\xRightarrow{\text{implicit fcn thm}}$ $L_x(x, \mu, \lambda) = 0$ has local solution $x(\mu, \lambda)$;

also, locally $L_{xx}(x, \mu, \lambda)$ pos. def. $\Rightarrow$ second order sufficient conditions for optimality fulfilled $\square$

Local duality

(becomes global duality for convex problems)

In the following we will assume all constraints to be active (else we have to ignore inactive constraints)

<u>Def</u>: The <u>(local) dual function</u> is defined for (μ,λ) near (μ^*,λ^*) as $\quad g(\mu,\lambda) = \min_{x \text{ near } x^*} L(x,\mu,\lambda)$.

Under (C) this is well-defined. The minimiser is denoted $x(\mu,\lambda)$.

<u>Thm</u>: $\quad Dg(\mu,\lambda) = \left(c_I(x(\mu,\lambda))^T \quad c_E(x(\mu,\lambda))^T \right)$

$$D^2 g(\mu,\lambda) = -\begin{pmatrix} Dc_I \\ Dc_E \end{pmatrix} L_{xx}(x(\mu,\lambda),\mu,\lambda)^{-1} \left(\nabla c_I \ \ \nabla c_E\right) \qquad \text{evaluated at } x = x(\mu,\lambda).$$

proof: Abbreviate $\bar x = x(\mu,\lambda)$, then $\quad g(\mu,\lambda) = f_0(\bar x) + \mu^T c_I(\bar x) + \lambda^T c_E(\bar x) = L(\bar x,\mu,\lambda)$

$$\Rightarrow Dg(\mu,\lambda) = \overbrace{L_x(\bar x,\mu,\lambda)}^{=0} D_{(\mu,\lambda)}\bar x + \left(c_I(\bar x)^T \quad c_E(\bar x)^T\right) = \left(c_I(\bar x)^T \quad c_E(\bar x)^T\right)$$

$$\Rightarrow D^2 g(\mu,\lambda) = \begin{pmatrix} Dc_I(\bar x) \\ Dc_E(\bar x) \end{pmatrix} D_{(\mu,\lambda)}\bar x$$

solve for $D_{(\mu,\lambda)}\bar x$ and plug in

Now differentiating $0 = L_x(\bar x,\mu,\lambda)$ wrt (μ,λ), $\quad 0 = L_{xx}(\bar x,\mu,\lambda) D_{(\mu,\lambda)}\bar x + \left(\nabla c_I(\bar x) \ \ \nabla c_E(\bar x)\right)$ $\square$

<u>Thm</u> (local duality): Let x^* be regular and locally optimal for (P) with Lagrange multipliers (μ^*,λ^*).

Under (C), the dual problem $\max_{(\mu,\lambda)\in\mathbb{R}^{m\times p}} g(\mu,\lambda)$ has the local solution (μ^*,λ^*) with $g(\mu^*,\lambda^*) = f_0(x^*)$ (strong duality)

proof: By the above, $x(\mu^*,\lambda^*) = x^*$ and $Dg(\mu^*,\lambda^*) = \left(c_I(x^*)^T \ \ c_E(x^*)^T\right) = 0$ & $D^2 g(\mu^*,\lambda^*)$ is neg. def.

$$\Rightarrow (\mu^*,\lambda^*) \text{ is local maximum of } g \text{ with } g(\mu^*,\lambda^*) = \min_{x \text{ near } x^*} f_0(x) + \mu^{*T} c_I(x) + \lambda^{*T} c_E(x) = f_0(x^*).$$ $\square$

Dual problems and penalty functions

- Steepest descent convergence rate for (P) depends on condition number κ of $L_{xx}(x^*,\mu^*,\lambda^*)$, restricted to the tangent subspace $T_{x^*}S$ to the active constraints.
- Steepest ascent convergence rate for dual problem depends on $\kappa(D^2 g(\mu^*,\lambda^*))$ with $D^2 g(\mu^*,\lambda^*) = -\begin{pmatrix} Dc_I \\ Dc_E \end{pmatrix} L_{xx}^{-1} (\nabla c_I \;\; \nabla c_E)$ at (x^*,μ^*,λ^*), a restriction of L_{xx}^{-1} to $T_{x^*}S^\perp$!

Problem: $L_{xx}(x^*,\mu^*,\lambda^*)$ is in general not pos. def. $\Rightarrow$ local duality not applicable!

- Steepest descent convergence rate of penalty method subproblem depends on condition number of $D^2 q(c_k, x_k)$, e.g. for $P(x_k) = \frac{1}{2}\|c_E(x_k)\|_2^2 + \frac{1}{2}\|c_I^+(x_k)\|_2^2$, $c_I^+ = \max(0, c_I)$

$$D^2 q(c_k,x_k) = D^2 f_0(x_k) + c_k\left[\sum_{i=1}^{p} g_i(x_k) D^2 g_i(x_k) + \sum_{i=1}^{m} f_i^+(x_k) D^2 f_i^+(x_k) + (\nabla c_I^+(x_k) \;\; \nabla c_E(x_k)) \begin{pmatrix} Dc_I^+(x_k) \\ Dc_E(x_k) \end{pmatrix}\right]$$

$$= L_{xx}(x_k,\mu_k,\lambda_k) + c_k (\nabla c_I^+(x_k) \;\; \nabla c_E(x_k)) \begin{pmatrix} Dc_I^+(x_k) \\ Dc_E(x_k) \end{pmatrix} \quad \text{for } (\mu_k,\lambda_k) = c_k(c_I^+(x_k), c_E(x_k))$$

Problem: As $c_k \to \infty$, the condition number gets arbitrarily bad (largest eigenvalue grows like c_k, smallest stays roughly same).

If Newton descent is used, then instead the efficiency of solving Newton's equations depends on condition number!

Augmented Lagrangian method: Motivation by duality

$$\min_{x\in\mathbb{R}^n} f_0(x) \text{ s.t. } c_E(x)=0 \iff \min_{x\in\mathbb{R}^n} f_0(x) + \frac{c}{2}\|c_E(x)\|_2^2 \text{ s.t. } c_E(x)=0$$

for any $c\in\mathbb{R}$!

Lagrangian of original formulation

- Lagrangian of new formulation is $L(x,\lambda) + \frac{c}{2}\|c_E(x)\|_2^2$.
- Same first order optimality condition and Lagrange multiplier λ^*, but Hessian at the optimum is $L_{xx}(x^*,\lambda^*) + c\nabla c_E(x^*)\nabla c_E(x^*)^T$, pos. def. for c large enough!

  $\Rightarrow$ dual method now applicable!
- Dual function $g(\lambda) = \min_{x \text{ near } x^*} f_0(x) + \lambda^T c_E(x) + \frac{c}{2}\|c_E(x)\|_2^2$ has at optimum the Hessian

  $$D^2 g(\lambda^*) = -Dc_E(x^*)\left[D^2 f_0(x^*) + \sum_{i=1}^p \lambda_i D^2 g_i(x^*) + c\nabla c_E(x^*)\nabla c_E(x^*)^T\right]^{-1}\nabla c_E(x^*)$$

  $\Rightarrow$ $D^2 g(\lambda^*)$ approaches $-\frac{1}{c}I$ for $c\to\infty$, which has condition number 1

  $\Rightarrow$ extremely fast convergence of dual ascent / very efficient Newton method
- Apply modified Newton's method to dual problem, using approximate Hessian $-\frac{1}{c}I$

  $$\Rightarrow \lambda_{k+1} = \lambda_k + c\nabla g(\lambda_k) = \lambda_k + c\, c_E(x_k)^T$$

  can also be viewed as gradient ascent with stepsize c

  where x_k minimises $f_0(x) + \lambda_k^T c_E(x_k) + \frac{c}{2}\|c_E(x_k)\|_2^2$

Augmented Lagrangian method: Motivation by penalty

$$\min_{x\in\mathbb{R}^n} f_0(x) \text{ s.t. } c_E(x)=0 \iff \min_{x\in\mathbb{R}^n} f_0(x) + \lambda^T c_E(x) \text{ s.t. } c_E(x)=0$$

for any $\lambda \in \mathbb{R}^p$

- penalty method objective $q(c,x) = f_0(x) + \lambda^T c_E(x) + \frac{c}{2}\|c_E(x)\|_2^2$

- if $\lambda = \lambda^*$ for the Lagrange multiplier of the original formulation, then the gradient of the penalty objective, $Df_0(x^*) + \lambda^T Dc_E(x^*) + c\, c_E(x^*)^T Dc_E(x^*)$, is zero

  $\Longrightarrow$ penalty method is exact even for finite c! $\Rightarrow$ Hessian better conditioned!

- optimal point satisfies $0 = Df_0(x^*) + \lambda^{*T} Dc_E(x^*)$

  penalty method iterate satisfies $0 = Df_0(x_k) + \left(\lambda_k + c\, c_E(x_k)\right)^T Dc_E(x_k)$

  $\Longrightarrow$ choose $\lambda_{k+1} \approx \lambda^* \approx \lambda_k + c\, c_E(x_k)$

Augmented Lagrangian method: algorithm

Def: $L_c(x,\lambda) = f_0(x) + \lambda^T c_E(x) + \frac{c}{2}\|c_E(x)\|_2^2$ is called the augmented Lagrangian to (P)

Alg: Given f_0, c_E, x_0, λ_0, monotonically increasing sequence $c_0, c_1, \dots$, $k=0$

repeat $\quad x_{k+1} = \operatorname{argmin}_{x\in\mathbb{R}^n} L_{c_k}(x,\lambda_k)$

$\quad\quad \lambda_{k+1} = \lambda_k + c_k\, c_E(x_{k+1})$

$\quad\quad k \leftarrow k+1$

until $c_E(x_{k+1}) = 0$

Thm: Let the 2nd order sufficient conditions for local optimality of (P) be satisfied at (x^*,λ^*).
There is a $c^*\in\mathbb{R}$ such that $L_c(x,\lambda^*)$ has a local minimum at x^* for all $c>c^*$.

proof:
- 1st order optimality satisfied at x^*, only need to check positive definiteness of $D_x^2 L_c$
- $D_x^2 L_c(x^*,\lambda^*) = L_{xx}(x^*,\lambda^*) + c\, Dc_E(x^*)^T Dc_E(x^*) = A + cB$

  A pos. def. on $T_{x^*}S = \ker B$
  B pos. def. on $T_{x^*}S^\perp = \operatorname{span}(\nabla c_E(x^*))$
- suppose, $\forall k\in\mathbb{N}\ \exists p_k : \|p_k\|_2^2 = 1,\ p_k^T(A+kB)p_k \le 0$, then for a subsequence, $p_k \to p$
- $0 \geq p_k^T B p_k \to p^T B p \implies p^T B p = 0 \implies p^T A p \le 0$, but this contradicts $p\in\ker B$ $\square$

Rem: By continuity, the above also holds in a neighbourhood of $(x^*,\lambda^*) \Rightarrow$ method well-defined

Augmented Lagrangian method: example

Ex: $\min\limits_{(x,y)\in\mathbb{R}^2} \quad 2x^2+2xy+y^2-2y \qquad \text{s.t.} \quad x=0$

- $L_c(x,y,\lambda) = 2x^2+2xy+y^2-2y + \lambda x + \frac{c}{2}x^2$
- $(x_{k+1},y_{k+1}) = \operatorname*{argmin}\limits_{(x,y)} L_{c_k}(x,y,\lambda_k) = \left(-\frac{2+\lambda_k}{2+c_k}, \frac{4+c_k+\lambda_k}{2+c_k}\right)$

  $\lambda_{k+1} = \lambda_k + c_k x_{k+1} = \frac{2}{2+c_k}\lambda_k - \frac{2c_k}{2+c_k}$
- $\lambda_k \xrightarrow[k\to\infty]{} -2$ with $(\lambda_{k+1}+2) = \frac{2}{2+c_k}(\lambda_k+2)$ $\Rightarrow$ linear convergence at rate $\frac{2}{2+c}$

Ex: $\min\limits_{(x,y)\in\mathbb{R}^2} \quad 2x^2+2xy+y^2-2y \qquad \text{s.t.} \quad x=0,\ y\leq 0$

- replace by $\min\limits_{(x,y,z)\in\mathbb{R}^2} \quad 2x^2+2xy+y^2-2y \qquad \text{s.t.} \quad x=0,\ y+z^2=0$
- $L_c(x,y,z,\lambda,\nu) = 2x^2+2xy+y^2-2y+\lambda x+\nu(y+z^2) + \frac{c}{2}x^2 + \frac{c}{2}(y+z^2)^2$
- $(x,y,z)_{k+1} = \left(\frac{2\nu_k-4-2\lambda_k-c_k\lambda_k}{(2+c_k)(4+c_k)-4}, \frac{2-\nu_k}{2+c_k} + \frac{2\lambda_k-(4\nu_k-8)/(2+c_k)}{(2+c_k)(4+c_k)-4}, 0\right)$

  $(\lambda,\nu)_{k+1} = (\lambda_k,\nu_k) + c_k(x_{k+1}, y_{k+1})$
- $(\lambda_k,\nu_k) \xrightarrow[k\to\infty]{} (0,2)$ (linear convergence for fixed c_k)

Augmented Lagrangian method: convergence rate

Recall: update $\lambda_{k+1} = \lambda_k + \overset{\text{assume fixed } c}{c}\, c_E(x_k) = \lambda_k + c\nabla g(\lambda_k)$ is gradient ascent

for dual function g to $\min\limits_{x \in \mathbb{R}^n} f_0(x) + \frac{c}{2}\|c_E(x)\|_2^2$ s.t. $c_E(x) = 0$

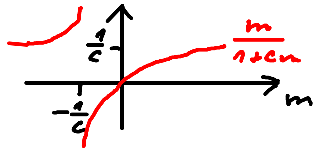
$\Rightarrow$ convergence rate depends on eigenvalues of

$$D^2 g(\lambda^*) = -\overbrace{Dc_E(x^*)}^{B}\Big[\overbrace{L_{xx}(x^*,\lambda^*)}^{A} + c\nabla c_E(x^*)\nabla c_E(x^*)^T\Big]^{-1}\nabla c_E(x^*) = -B(A+cB^TB)^{-1}B^T \quad \leftarrow \text{Hessian of dual problem to (P)}$$

Thm: Let m be an eigenvalue of $BA^{-1}B^T$, then $\frac{m}{1+cm}$ is an eigenvalue of $B(A+cB^TB)^{-1}B^T$.

proof: • By Sherman-Morrison-Woodbury, $B(A+cB^TB)^{-1}B^T = F - c\,F(I+cF)^{-1}F$ for $F = BA^{-1}B^T$

• let $Fv = mv$, then $B(A+cB^TB)^{-1}B^T v = \left(m - c\,\frac{m^2}{1+cm}\right)v = \frac{m}{1+cm}v$ $\square$



Rem: • The convergence rate of steepest ascent for the dual problem to (P) depends on the condition number $\kappa(BA^{-1}B^T) = \frac{M}{m}$ for smallest / largest eigenvalue m / M of $BA^{-1}B^T$.

• Convergence rate of augmented Lagrangian method depends on $\kappa(B(A+cB^TB)^{-1}B^T) = \dfrac{c+\frac{1}{m}}{c+\frac{1}{M}}$.

• One can improve on the rate by using a quasi-Newton or Newton method for g.