

Continuous optimization

ENT305A

Laurent Pfeiffer

Inria and CentraleSupélec, Université Paris-Saclay

Ensta-Paris
Institut Polytechnique de Paris
September 2023



Organisation

Organization:

- Class 1: lecture
- Class 2: lecture (1h 30) + programming exercises (2h)
- Class 3: lecture (1h 30) + programming exercises (2h)
- Class 4: programming exercises
- Class 5: programming exercise (1h 30) + exam (2h).

Contact me:

laurent.pfeiffer@inria.fr

Main objectives

Skills to be developped:

- **Modelling** of practical situations involving decision variables, constraints to be satisfied, a cost to minimize, as an *optimization problem*.
- **Numerical resolution** of such problems with the help of AMPL (A Mathematical Programming Language).
- Basic knowledge in optimization **theory and numerics**.

Pre-requisite:

- Programming: none
- Maths: little.

100

1 General introduction

- What is an optimization problem?
- Classes of problems
- Existence of a solution
- Derivatives

2 Methods for unconstrained optimization

- Optimality conditions
- Gradient methods
- Newton's method

3 Optimality conditions for constrained problems

- Linear constraints
- Non-linear constraints
- Sensitivity analysis

What is an optimization problem?

Definition 1

An **optimization problem** is a mathematical expression of the form:

$$\inf_{x \in \mathcal{D}} f(x), \quad \text{subject to: } x \in K, \quad (P)$$

where:

- \mathcal{D} is a set, called **domain** of f
- $f: \mathcal{D} \rightarrow \mathbb{R}$ is called **cost function** (or **objective function**)
- $K \subseteq \mathcal{D}$ is called **feasible set**.

In this class: $\mathcal{D} = \mathbb{R}^n$. **Unconstrained** optimization: $\mathcal{D} = K = \mathbb{R}^n$.

Straightforward adaptation of all results of the class to **maximization** problems, replacing f by $-f$.

Abbreviation: “subject to” \rightsquigarrow “s.t.”.

1. $\frac{1}{2}$ 2. $\frac{1}{3}$ 3. $\frac{1}{4}$ 4. $\frac{1}{5}$ 5. $\frac{1}{6}$ 6. $\frac{1}{7}$ 7. $\frac{1}{8}$ 8. $\frac{1}{9}$ 9. $\frac{1}{10}$ 10. $\frac{1}{11}$ 11. $\frac{1}{12}$ 12. $\frac{1}{13}$ 13. $\frac{1}{14}$ 14. $\frac{1}{15}$ 15. $\frac{1}{16}$ 16. $\frac{1}{17}$ 17. $\frac{1}{18}$ 18. $\frac{1}{19}$ 19. $\frac{1}{20}$ 20. $\frac{1}{21}$ 21. $\frac{1}{22}$ 22. $\frac{1}{23}$ 23. $\frac{1}{24}$ 24. $\frac{1}{25}$ 25. $\frac{1}{26}$ 26. $\frac{1}{27}$ 27. $\frac{1}{28}$ 28. $\frac{1}{29}$ 29. $\frac{1}{30}$ 30. $\frac{1}{31}$ 31. $\frac{1}{32}$ 32. $\frac{1}{33}$ 33. $\frac{1}{34}$ 34. $\frac{1}{35}$ 35. $\frac{1}{36}$ 36. $\frac{1}{37}$ 37. $\frac{1}{38}$ 38. $\frac{1}{39}$ 39. $\frac{1}{40}$ 40. $\frac{1}{41}$ 41. $\frac{1}{42}$ 42. $\frac{1}{43}$ 43. $\frac{1}{44}$ 44. $\frac{1}{45}$ 45. $\frac{1}{46}$ 46. $\frac{1}{47}$ 47. $\frac{1}{48}$ 48. $\frac{1}{49}$ 49. $\frac{1}{50}$ 50. $\frac{1}{51}$ 51. $\frac{1}{52}$ 52. $\frac{1}{53}$ 53. $\frac{1}{54}$ 54. $\frac{1}{55}$ 55. $\frac{1}{56}$ 56. $\frac{1}{57}$ 57. $\frac{1}{58}$ 58. $\frac{1}{59}$ 59. $\frac{1}{60}$ 60. $\frac{1}{61}$ 61. $\frac{1}{62}$ 62. $\frac{1}{63}$ 63. $\frac{1}{64}$ 64. $\frac{1}{65}$ 65. $\frac{1}{66}$ 66. $\frac{1}{67}$ 67. $\frac{1}{68}$ 68. $\frac{1}{69}$ 69. $\frac{1}{70}$ 70. $\frac{1}{71}$ 71. $\frac{1}{72}$ 72. $\frac{1}{73}$ 73. $\frac{1}{74}$ 74. $\frac{1}{75}$ 75. $\frac{1}{76}$ 76. $\frac{1}{77}$ 77. $\frac{1}{78}$ 78. $\frac{1}{79}$ 79. $\frac{1}{80}$ 80. $\frac{1}{81}$ 81. $\frac{1}{82}$ 82. $\frac{1}{83}$ 83. $\frac{1}{84}$ 84. $\frac{1}{85}$ 85. $\frac{1}{86}$ 86. $\frac{1}{87}$ 87. $\frac{1}{88}$ 88. $\frac{1}{89}$ 89. $\frac{1}{90}$ 90. $\frac{1}{91}$ 91. $\frac{1}{92}$ 92. $\frac{1}{93}$ 93. $\frac{1}{94}$ 94. $\frac{1}{95}$ 95. $\frac{1}{96}$ 96. $\frac{1}{97}$ 97. $\frac{1}{98}$ 98. $\frac{1}{99}$ 99. $\frac{1}{100}$ 100. $\frac{1}{101}$ 101. $\frac{1}{102}$ 102. $\frac{1}{103}$ 103. $\frac{1}{104}$ 104. $\frac{1}{105}$ 105. $\frac{1}{106}$ 106. $\frac{1}{107}$ 107. $\frac{1}{108}$ 108. $\frac{1}{109}$ 109. $\frac{1}{110}$ 110. $\frac{1}{111}$ 111. $\frac{1}{112}$ 112. $\frac{1}{113}$ 113. $\frac{1}{114}$ 114. $\frac{1}{115}$ 115. $\frac{1}{116}$ 116. $\frac{1}{117}$ 117. $\frac{1}{118}$ 118. $\frac{1}{119}$ 119. $\frac{1}{120}$ 120. $\frac{1}{121}$ 121. $\frac{1}{122}$ 122. $\frac{1}{123}$ 123. $\frac{1}{124}$ 124. $\frac{1}{125}$ 125. $\frac{1}{126}$ 126. $\frac{1}{127}$ 127. $\frac{1}{128}$ 128. $\frac{1}{129}$ 129. $\frac{1}{130}$ 130. $\frac{1}{131}$ 131. $\frac{1}{132}$ 132. $\frac{1}{133}$ 133. $\frac{1}{134}$ 134. $\frac{1}{135}$ 135. $\frac{1}{136}$ 136. $\frac{1}{137}$ 137. $\frac{1}{138}$ 138. $\frac{1}{139}$ 139. $\frac{1}{140}$ 140. $\frac{1}{141}$ 141. $\frac{1}{142}$ 142. $\frac{1}{143}$ 143. $\frac{1}{144}$ 144. $\frac{1}{145}$ 145. $\frac{1}{146}$ 146. $\frac{1}{147}$ 147. $\frac{1}{148}$ 148. $\frac{1}{149}$ 149. $\frac{1}{150}$ 150. $\frac{1}{151}$ 151. $\frac{1}{152}$ 152. $\frac{1}{153}$ 153. $\frac{1}{154}$ 154. $\frac{1}{155}$ 155. $\frac{1}{156}$ 156. $\frac{1}{157}$ 157. $\frac{1}{158}$ 158. $\frac{1}{159}$ 159. $\frac{1}{160}$ 160. $\frac{1}{161}$ 161. $\frac{1}{162}$ 162. $\frac{1}{163}$ 163. $\frac{1}{164}$ 164. $\frac{1}{165}$ 165. $\frac{1}{166}$ 166. $\frac{1}{167}$ 167. $\frac{1}{168}$ 168. $\frac{1}{169}$ 169. $\frac{1}{170}$ 170. $\frac{1}{171}$ 171. $\frac{1}{172}$ 172. $\frac{1}{173}$ 173. $\frac{1}{174}$ 174. $\frac{1}{175}$ 175. $\frac{1}{176}$ 176. $\frac{1}{177}$ 177. $\frac{1}{178}$ 178. $\frac{1}{179}$ 179. $\frac{1}{180}$ 180. $\frac{1}{181}$ 181. $\frac{1}{182}$ 182. $\frac{1}{183}$ 183. $\frac{1}{184}$ 184. $\frac{1}{185}$ 185. $\frac{1}{186}$ 186. $\frac{1}{187}$ 187. $\frac{1}{188}$ 188. $\frac{1}{189}$ 189. $\frac{1}{190}$ 190. $\frac{1}{191}$ 191. $\frac{1}{192}$ 192. $\frac{1}{193}$ 193. $\frac{1}{194}$ 194. $\frac{1}{195}$ 195. $\frac{1}{196}$ 196. $\frac{1}{197}$ 197. $\frac{1}{198}$ 198. $\frac{1}{199}$ 199. $\frac{1}{200}$ 200. $\frac{1}{201}$ 201. $\frac{1}{202}$ 202. $\frac{1}{203}$ 203. $\frac{1}{204}$ 204. $\frac{1}{205}$ 205. $\frac{1}{206}$ 206. $\frac{1}{207}$ 207. $\frac{1}{208}$ 208. $\frac{1}{209}$ 209. $\frac{1}{210}$ 210. $\frac{1}{211}$ 211. $\frac{1}{212}$ 212. $\frac{1}{213}$ 213. $\frac{1}{214}$ 214. $\frac{1}{215}$ 215. $\frac{1}{216}$ 216. $\frac{1}{217}$ 217. $\frac{1}{218}$ 218. $\frac{1}{219}$ 219. $\frac{1}{220}$ 220. $\frac{1}{221}$ 221. $\frac{1}{222}$ 222. $\frac{1}{223}$ 223. $\frac{1}{224}$ 224. $\frac{1}{225}$ 225. $\frac{1}{226}$ 226. $\frac{1}{227}$ 227. $\frac{1}{228}$ 228. $\frac{1}{229}$ 229. $\frac{1}{230}$ 230. $\frac{1}{231}$ 231. $\frac{1}{232}$ 232. $\frac{1}{233}$ 233. $\frac{1}{234}$ 234. $\frac{1}{235}$ 235. $\frac{1}{236}$ 236. $\frac{1}{237}$ 237. $\frac{1}{238}$ 238. $\frac{1}{239}$ 239. $\frac{1}{240}$ 240.

(1) $\frac{1}{2} \leq \frac{1}{2} - \frac{1}{2} = 0$

-

What is an optimization problem?

Remarks.

- An optimization problem may **not** have a solution. *Examples:*

$$\inf_{x \in \mathbb{R}} e^x, \quad (P_1)$$

$$\inf_{x \in \mathbb{R}} x^3. \quad (P_2)$$

- The concept of **value** of an optimization can also be defined whether the problem has a solution or not, as an element of $\mathbb{R} \cup \{-\infty, \infty\}$. In particular:

$$\text{val}(P_1) = 0, \quad \text{val}(P_2) = -\infty.$$

What is an optimization problem?

Definition 3

Let $\bar{x} \in K$. We call \bar{x} a **local solution** to (P) if there exists $\varepsilon > 0$ such that the following holds true: for all $x \in K$,

$$\|x - \bar{x}\| \leq \varepsilon \implies f(x) \geq f(\bar{x}).$$

Example: $\inf_{x \in \mathbb{R}} -x^2$, s.t. $x \in [-1, 2]$. Local solutions: -1 and 2 .

Remarks.

- A global solution is also a local solution.
- The notion of local optimality does not depend on the norm, if K is a subset of a finite dimensional vector space.

1. *Journal of the American Medical Association*, 2000; 283: 2689-2696.

—

1. *Chlorophyll a* and *Chlorophyll b* were determined by the method of Arar and Collins (1987).

A feasible point \bar{x} is a local solution to (P) if and only if there

$$\inf_{x \in \mathbb{R}^n} f(x), \quad x \in K \cap \bar{B}(\bar{x}, \varepsilon).$$

What is an optimization problem?

Constraints.

Most of the time, the feasible set K is described by

$$K = \left\{ x \in \mathbb{R}^n \mid \begin{array}{l} g_i(x) \geq 0, \quad \forall i \in \mathcal{I} \\ g_i(x) = 0, \quad \forall i \in \mathcal{E} \end{array} \right\},$$

where $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $(\mathcal{I}, \mathcal{E})$ is a partition of $\{1, \dots, m\}$.

We call the expressions

- $g_i(x) \geq 0$: **inequality constraint**
- $g_i(x) = 0$: **equality constraint.**

1 General introduction

- What is an optimization problem?
- **Classes of problems**
- Existence of a solution
- Derivatives

2 Methods for unconstrained optimization

- Optimality conditions
- Gradient methods
- Newton's method

3 Optimality conditions for constrained problems

- Linear constraints
- Non-linear constraints
- Sensitivity analysis

1. *Journal of Management Studies*, 1996, 33, 1, 1-15.

- 1 Economical problems
- 2 Physical problems
- 3 Inverse problems
- 4 Learning problems.

Any practical situation involving

- ◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ 🔍 ↺

1. *Journal of the American Medical Association*, 2000; 284: 2689-2695.

- Mechanical structures
- Electricity networks
- Gas networks

- Cournot models with competing firms
- Traffic models on road networks.

1. *Journal of Management Studies*, 1997, 34, 1, 1-15.

- the epicenter x of an earthquake, given seismic measurements y .
- localization x of a crack in a mechanical structure, given displacements measurements y provided by captors
- temperature in the core of nuclear plant, given external temperature measurements

1. *Journal of the American Medical Association*, 2000; 283: 2689-2696.

- may not have a solution (because of inaccurate measurements)
- may have several solutions (too few measurements).

$$\inf_{x \in \mathcal{D}} \|y - F(x)\|^2, \quad \text{subject to: } x \in K,$$

where the constraints may model a priori knowledge on x .

1. *Journal of Management Studies*, 1997, 34, 1, 1-14.

(5)

Classes of optimization problems

Approach by least-square.

- Fix an explicit function $F: X \times R \rightarrow Y$, for some parameter set R .
- “Tune the parameter r ”, that is, solve the least-squares problem

$$\inf_{r \in R} \sum_{k=1}^K \|y_k - F(x_k, r)\|^2.$$

- Heuristical relation $y \approx F(x, \bar{r})$ for the optimal \bar{r} .

Example: linear regression. For $X = \mathbb{R}$ and $Y = \mathbb{R}$, take $R = \mathbb{R}^2$ and

$$F(x, r) = r_1 + r_2 x.$$

1 General introduction

- What is an optimization problem?
- Classes of problems
- **Existence of a solution**
- Derivatives

2 Methods for unconstrained optimization

- Optimality conditions
- Gradient methods
- Newton's method

3 Optimality conditions for constrained problems

- Linear constraints
- Non-linear constraints
- Sensitivity analysis

$$\|x_k\| \rightarrow \infty \implies f(x_k) \rightarrow +\infty.$$

Remark. The definition is independent of the norm.

Downloaded from <http://ajph.org/> on November 10, 2014

Existence of a solution

Solution. We have $x^4 \geq 2x^2 - 1$, since

$$0 \leq (x^2 - 1)^2 = x^4 - 2x^2 + 1.$$

Therefore

$$\begin{aligned} f(x, y) &\geq 2x^2 - 1 - 2xy + 2y^2 \\ &= (x^2 + y^2) - 1 + (x - y)^2 \\ &\geq \|(x, y)\|^2 - 1 \xrightarrow{\|(x, y)\| \rightarrow \infty} \infty, \end{aligned}$$

where $\|\cdot\|$ denotes the Euclidean norm. Thus f is coercive.

Assume the following:

- K is non-empty and closed
- f is continuous on K
- f is coercive on K .

Then the optimization problem (P) has (at least) one solution.

Remarks.

- If $K = \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, \forall i \in \mathcal{I}, g_i(x) = 0, \forall i \in \mathcal{E}\}$, where g is continuous, then K is closed.
- If K is bounded, then f is coercive because there is no sequence $(x_k)_{k \in \mathbb{N}}$ in K such that $\|x_k\| \rightarrow \infty$.

Existence of a solution

Elements of proof. Fix $x_0 \in K$ and define:

$$K' = \{x \in K \mid f(x) \leq f(x_0)\}.$$

- K' is **closed**, since f is continuous and K is closed
- K' is **bounded**, since f is coercive (proof by contradiction)
- The problem $\inf f(x)$, s.t. $x \in K'$ is known to have a **global solution** \bar{x} since f is continuous and K' compact.
- Show that \bar{x} is a global solution to the original problem.

1 General introduction

- What is an optimization problem?
- Classes of problems
- Existence of a solution
- **Derivatives**

2 Methods for unconstrained optimization

- Optimality conditions
- Gradient methods
- Newton's method

3 Optimality conditions for constrained problems

- Linear constraints
- Non-linear constraints
- Sensitivity analysis

Derivatives

Definition 6

A function $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is called **differentiable** at \bar{x} if for all $i = 1, \dots, m$, for all $j = 1, \dots, n$, the function

$$x \in \mathbb{R} \mapsto F_i(\bar{x}_1, \dots, \bar{x}_{j-1}, x, \bar{x}_{j+1}, \dots) \in \mathbb{R}$$

is differentiable. Its derivative at \bar{x}_j is called **partial derivative** of F , it is denoted $\frac{\partial F_i}{\partial x_j}(\bar{x})$.

The matrix

$$DF(\bar{x}) = \left(\frac{\partial F_i}{\partial x_j}(\bar{x}) \right)_{\substack{i=1, \dots, m \\ j=1, \dots, n}} \in \mathbb{R}^{m \times n}$$

is called **Jacobian** matrix.

- $$F(x + \delta x) = F(x) + DF(x)\delta x + o(\|\delta x\|).$$

(1) $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2$, where $\mathcal{A}_1 = \{A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9, A_{10}, A_{11}, A_{12}, A_{13}, A_{14}, A_{15}, A_{16}, A_{17}, A_{18}, A_{19}, A_{20}, A_{21}, A_{22}, A_{23}, A_{24}, A_{25}, A_{26}, A_{27}, A_{28}, A_{29}, A_{30}, A_{31}, A_{32}, A_{33}, A_{34}, A_{35}, A_{36}, A_{37}, A_{38}, A_{39}, A_{40}, A_{41}, A_{42}, A_{43}, A_{44}, A_{45}, A_{46}, A_{47}, A_{48}, A_{49}, A_{50}, A_{51}, A_{52}, A_{53}, A_{54}, A_{55}, A_{56}, A_{57}, A_{58}, A_{59}, A_{60}, A_{61}, A_{62}, A_{63}, A_{64}, A_{65}, A_{66}, A_{67}, A_{68}, A_{69}, A_{70}, A_{71}, A_{72}, A_{73}, A_{74}, A_{75}, A_{76}, A_{77}, A_{78}, A_{79}, A_{80}, A_{81}, A_{82}, A_{83}, A_{84}, A_{85}, A_{86}, A_{87}, A_{88}, A_{89}, A_{90}, A_{91}, A_{92}, A_{93}, A_{94}, A_{95}, A_{96}, A_{97}, A_{98}, A_{99}, A_{100}\}$ and $\mathcal{A}_2 = \{A_{101}, A_{102}, A_{103}, A_{104}, A_{105}, A_{106}, A_{107}, A_{108}, A_{109}, A_{110}, A_{111}, A_{112}, A_{113}, A_{114}, A_{115}, A_{116}, A_{117}, A_{118}, A_{119}, A_{120}, A_{121}, A_{122}, A_{123}, A_{124}, A_{125}, A_{126}, A_{127}, A_{128}, A_{129}, A_{130}, A_{131}, A_{132}, A_{133}, A_{134}, A_{135}, A_{136}, A_{137}, A_{138}, A_{139}, A_{140}, A_{141}, A_{142}, A_{143}, A_{144}, A_{145}, A_{146}, A_{147}, A_{148}, A_{149}, A_{150}, A_{151}, A_{152}, A_{153}, A_{154}, A_{155}, A_{156}, A_{157}, A_{158}, A_{159}, A_{160}, A_{161}, A_{162}, A_{163}, A_{164}, A_{165}, A_{166}, A_{167}, A_{168}, A_{169}, A_{170}, A_{171}, A_{172}, A_{173}, A_{174}, A_{175}, A_{176}, A_{177}, A_{178}, A_{179}, A_{180}, A_{181}, A_{182}, A_{183}, A_{184}, A_{185}, A_{186}, A_{187}, A_{188}, A_{189}, A_{190}, A_{191}, A_{192}, A_{193}, A_{194}, A_{195}, A_{196}, A_{197}, A_{198}, A_{199}, A_{200}\}$.

- $$DH(x) = DF(G(x))DG(x), \quad \text{for all } x \in \mathbb{R}^p.$$

() () () ()

Derivatives

Definition 7

Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ be differentiable at $x \in \mathbb{R}^n$. We call **gradient** of f (at x) the column vector:

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix} = Df(x)^\top.$$

Derivatives

Definition 8

The function $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be **twice differentiable** if it is differentiable and DF is differentiable.

We denote: $\frac{\partial^2 F_i}{\partial x_j \partial x_k}(x) = \frac{\partial}{\partial x_j} \left(\frac{\partial F}{\partial x_k} \right)(x)$.

If $m = 1$, the matrix

$$D^2 F(x) = \left(\frac{\partial^2 F}{\partial x_j \partial x_k}(x) \right)_{\substack{j=1,\dots,n \\ k=1,\dots,n}}$$

is called **Hessian** matrix. It is symmetric.

Derivatives

Exercise.

Calculate the gradient and the Hessian of the function

$$f: x \in \mathbb{R}^n \mapsto \frac{1}{2} \langle x, Ax \rangle + \langle b, x \rangle,$$

where $A \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$.

Derivatives

Solution. We have

$$f(x) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n A_{ij} x_i x_j + \sum_{i=1}^n b_i x_i.$$

Therefore,

$$\begin{aligned} \frac{\partial f}{\partial x_k}(x) &= \frac{1}{2} \sum_{j=1}^n A_{kj} x_j + \frac{1}{2} \sum_{i=1}^n A_{ik} x_i + b_k \\ &= \frac{1}{2} (Ax)_k + \frac{1}{2} (A^\top x)_k + b_k. \end{aligned}$$

Therefore,

$$\nabla f(x) = \frac{1}{2} (A + A^\top) x + b.$$

Hessian: $D^2 f(x) = \frac{1}{2} (A + A^\top).$

1 General introduction

- What is an optimization problem?
- Classes of problems
- Existence of a solution
- Derivatives

2 Methods for unconstrained optimization

- Optimality conditions
- Gradient methods
- Newton's method

3 Optimality conditions for constrained problems

- Linear constraints
- Non-linear constraints
- Sensitivity analysis

Optimality conditions

Let us fix a continuously differentiable function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ for the whole section. Let us consider

$$\inf_{x \in \mathbb{R}^n} f(x) \quad (P)$$

The function f is said to be **stationary** at $x \in \mathbb{R}^n$ if $\nabla f(x) = 0$.

Theorem 9 (Necessary optimality condition)

Let $\bar{x} \in \mathbb{R}^n$ be a local solution of (P). Then, f is stationary at \bar{x} .

Remark. Stationarity is only a necessary condition!

Optimality conditions

Theorem 10

Assume that f is twice continuously differentiable. Let \bar{x} be a stationary point.

■ *Necessary condition.*

If \bar{x} is a local solution of (P) , then $D^2f(\bar{x})$ is positive semi-definite, that is to say,

$$\langle h, D^2f(\bar{x})h \rangle \geq 0, \quad \text{for all } h \in \mathbb{R}^n.$$

■ *Sufficient condition.*

If $D^2f(\bar{x})$ is positive definite, that is to say if

$$\langle h, D^2f(\bar{x})h \rangle > 0, \quad \text{for all } h \in \mathbb{R}^n \setminus \{0\},$$

then \bar{x} is a local solution of (P) .

Optimality conditions

Definition 11

The function f is said to be convex if

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y),$$

for all x and $y \in \mathbb{R}^n$ and for all $\lambda \in [0, 1]$.

Theorem 12

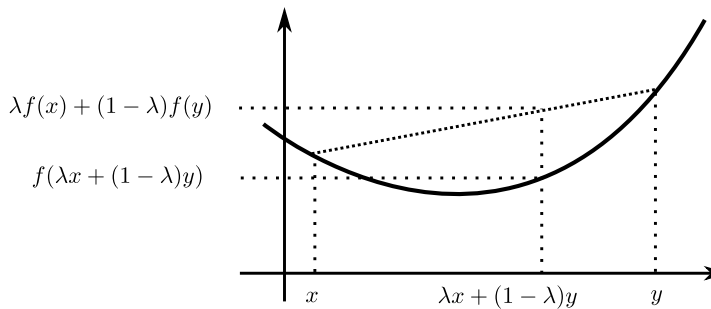
- *The function f is convex if and only if*

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle,$$

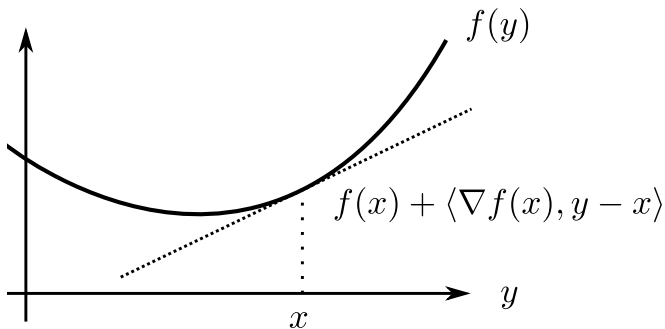
for all x and $y \in \mathbb{R}^n$.

- *If f is twice differentiable, then f is convex if and only if $D^2 f(x)$ is symmetric positive semi-definite for all $x \in \mathbb{R}^n$.*

Optimality conditions



Optimality conditions



Optimality conditions

Theorem 13

Assume that f is convex. Let \bar{x} be a stationary point of f . Then it is a global solution of (P) .

Proof. For all $x \in \mathbb{R}^n$, we have

$$f(x) \geq f(\bar{x}) + \langle \nabla f(\bar{x}), x - \bar{x} \rangle = f(\bar{x}).$$

Optimality conditions

Exercise.

Let $A \in \mathbb{R}^{n \times n}$ be symmetric positive definite and let $b \in \mathbb{R}^n$. Let

$$f : x \in \mathbb{R}^n \mapsto \frac{1}{2} \langle x, Ax \rangle + \langle b, x \rangle.$$

Prove that

$$\inf_{x \in \mathbb{R}^n} f(x)$$

has a unique solution. *Hint:* A is regular.

Optimality conditions

Solution.

- We have

$$\nabla f(x) = Ax + b.$$

Since A is regular, there exists a unique solution \bar{x} to the equation $Ax + b = 0$ (that is, $\bar{x} = -A^{-1}b$).

- We also have $D^2f(x) = A$. Therefore, f is convex and \bar{x} is a global solution.
- Any solution is stationary. Since \bar{x} is the unique stationary point, it is also the unique solution.

1 General introduction

- What is an optimization problem?
- Classes of problems
- Existence of a solution
- Derivatives

2 Methods for unconstrained optimization

- Optimality conditions
- **Gradient methods**
- Newton's method

3 Optimality conditions for constrained problems

- Linear constraints
- Non-linear constraints
- Sensitivity analysis

0 0 0 0

$$\inf_{x \in \mathbb{R}^n} f(x). \quad (P)$$

$$f(x_{k+1}) \leq f(x_k), \quad \forall k \in \mathbb{N},$$

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

$$\inf_{x \in \mathbb{R}^n} f(x). \quad (P)$$

$$f(x_{k+1}) \leq f(x_k), \quad \forall k \in \mathbb{N},$$

Gradient methods

Definition 14

Let $x \in \mathbb{R}^n$ and let $d \in \mathbb{R}^n$. The vector d is called **descent direction** if

$$\langle \nabla f(x), d \rangle < 0.$$

Remark. If $\nabla f(x) \neq 0$, then $d = -\nabla f(x)$ is a descent direction. Indeed,

$$\langle \nabla f(x), d \rangle = -\langle \nabla f(x), \nabla f(x) \rangle = -\|\nabla f(x)\|^2 < 0.$$

Gradient methods

Main idea of gradient methods.

Let $x_k \in \mathbb{R}^n$. Let d_k be a descent direction at x_k . Let $\alpha > 0$. Then

$$f(x_k + \alpha d_k) = f(x_k) + \underbrace{\alpha \langle \nabla f(x_k), d_k \rangle}_{<0} + o(\alpha).$$

Therefore, if α is small enough,

$$f(x_k + \alpha d_k) < f(x_k).$$

We can set

$$x_{k+1} = x_k + \alpha d_k.$$

Gradient methods

On the choice of α_k .

Let us fix $x_k \in \mathbb{R}^n$. Let us define

$$\phi_k: \alpha \in \mathbb{R} \mapsto f(x_k + \alpha d_k).$$

The condition $f(x_k + \alpha_k d_k) < f(x_k)$ is equivalent to

$$\phi_k(\alpha_k) < \phi_k(0).$$

A natural idea: define α_k as a solution to

$$\inf_{\alpha \geq 0} \phi_k(\alpha).$$

Minimizing ϕ_k would take too much time! A **compromise** must be found between simplicity of computation and quality of α .

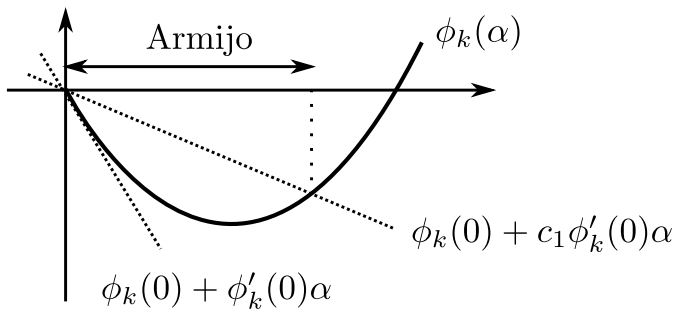
1. $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

1. $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

100

[illegible]

Gradient methods



Gradient methods

Backstepping algorithm for Armijo's rule

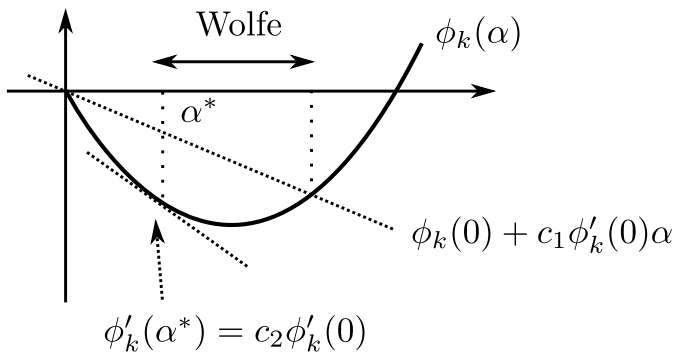
- 1 Input: $c_1 \in (0, 1)$, $\beta > 0$, and $\gamma \in (0, 1)$.
- 2 Set $\alpha = \beta$.
- 3 While α does not satisfy Armijo's rule,
 - Set $\alpha = \gamma\alpha$.
- 4 Output α .

Definition 16

Let $0 < c_1 < c_2 < 1$. We say that $\alpha > 0$ satisfies **Wolfe's rule** if

$$\phi_k(\alpha) < \phi_k(0) + c_1 \phi'_k(0) \alpha \quad \text{and} \quad \phi'_k(\alpha) \geq c_2 \phi'(0).$$

Gradient methods



Gradient methods

Bisection method for Wolfe's rule

- 1 Input: $c_1 \in (0, 1)$, $c_2 \in (c_1, 1)$, $\beta > 0$, $\eta > 1$.
- 2 Set $\alpha = \beta$.
- 3 While α satisfies Armijo's rule, do
 - Set $\alpha = \eta\alpha$.
- 4 Set $\alpha_{\min} = 0$, α_{\max} , $\alpha = \frac{1}{2}(\alpha_{\min} + \alpha_{\max})$.
- 5 While α does not satisfy Wolfe's rule, do
 - If α does not satisfy Armijo's rule, then set $\alpha_{\max} = \alpha$.
Else set $\alpha_{\min} = \alpha$.
 - Set $\alpha = \frac{1}{2}(\alpha_{\min} + \alpha_{\max})$.
- 6 Output: α .

General comments on the direct effect of the risk factors

- The algorithms for the computation of stepsizes satisfying Armijo and Wolfe's rules converge in **finitely many iterations** (under non-restrictive assumptions).
- Without convexity assumption on f , very little can be said about the convergence of the sequence $(x_k)_{k \in \mathbb{N}}$. Typical results ensure that any accumulation point is stationary.
- In practice: $(x_k)_{k \in \mathbb{N}}$ “usually” **converges to a local solution**. Thus a good **initialization** (that is the choice of x_0) is crucial.
- In general, **slow** convergence.

1 General introduction

- What is an optimization problem?
- Classes of problems
- Existence of a solution
- Derivatives

2 Methods for unconstrained optimization

- Optimality conditions
- Gradient methods
- **Newton's method**

3 Optimality conditions for constrained problems

- Linear constraints
- Non-linear constraints
- Sensitivity analysis

Originally, Newton's method aims at solving non-linear equations of the form

where $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a given continuously differentiable function. It is an iterative method, generating a sequence $(x_k)_{k \in \mathbb{N}}$. Given x_k , we have

Thus we look x_{k+1} as the solution to the linear equation

that is, $x_{k+1} = x_k - DF(x_k)^{-1}F(x_k)$.

Newton's method

Remarks.

- If there exists \bar{x} such that $F(\bar{x}) = 0$ and $DF(\bar{x})$ is regular, then for x_0 close enough to \bar{x} , the sequence $(x_k)_{k \in \mathbb{N}}$ is well-posed converges “**quickly**” to \bar{x} .
- On the other hand, if x_0 is far away from \bar{x} , there is **no guaranty** of convergence.

Back to problem (P). Assume that f is continuously twice differentiable. Apply Newton's method with $F(x) = \nabla f(x)$ so as to solve $\nabla f(x) = 0$. Update formula:

$$x_{k+1} = x_k - D^2 f(x_k)^{-1} \nabla f(x_k).$$

The difficulties mentioned above are still relevant.

Newton's method

Globalization of Newton's method.

- Newton's formula can be written in the form:

$$x_{k+1} = x_k + \alpha_k d_k,$$

where

$$\alpha_k = 1 \quad \text{and} \quad d_k = -D^2 f(x_k)^{-1} \nabla f(x_k).$$

- If $D^2 f(x_k)$ is positive definite (and $\nabla f(x_k) \neq 0$), then $D^2 f(x_k)^{-1}$ is also positive definite, and therefore d_k is descent direction:

$$\langle \nabla f(x_k), d_k \rangle = -\langle \nabla f(x_k), D^2 f(x_k)^{-1} \nabla f(x_k) \rangle < 0.$$

Newton's method

Globalised Newton's method.

- 1 Input: $x_0 \in \mathbb{R}^n$, $\varepsilon > 0$, a linesearch rule (Armijo, Wolfe,...).
- 2 Set $k = 0$.
- 3 While $\|\nabla f(x_k)\| \geq \varepsilon$, do
 - (a) If $-D^2f(x_k)^{-1}\nabla f(x_k)$ is computable and is a descent direction, set $d_k = -D^2f(x_k)^{-1}\nabla f(x_k)$, otherwise set $d_k = -\nabla f(x_k)$.
 - (b) If $\alpha = 1$ satisfies the linesearch rule, then set $\alpha_k = 1$. Otherwise, find α_k with an appropriate method.
 - (c) Set $x_{k+1} = x_k + \alpha_k d_k$.
 - (d) Set $k = k + 1$.
- 4 Output: x_k .

1 General introduction

- What is an optimization problem?
- Classes of problems
- Existence of a solution
- Derivatives

2 Methods for unconstrained optimization

- Optimality conditions
- Gradient methods
- Newton's method

3 Optimality conditions for constrained problems

- Linear constraints
- Non-linear constraints
- Sensitivity analysis

1. 9

- $\frac{1}{\sqrt{2}}$

Linear equality constraints

Theorem 17

Assume that g is affine, that is to say, there exists $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ such that

$$g(x) = Ax + b.$$

Let \bar{x} be a local solution to (P) .

Then there exists $\lambda \in \mathbb{R}^m$ such that the following three conditions, referred to as Karush-Kuhn-Tucker (KKT) conditions, are satisfied:

- 1 Stationarity condition: $\nabla_x L(\bar{x}, \lambda) = 0$.
- 2 Sign condition: for all $i \in \mathcal{I}$, $\lambda_i \geq 0$.
- 3 Complementarity condition: for all $i \in \mathcal{I}$,
 $g_i(\bar{x}) > 0 \implies \lambda_i = 0$.

1. 9

Linear constraints

Exercise.

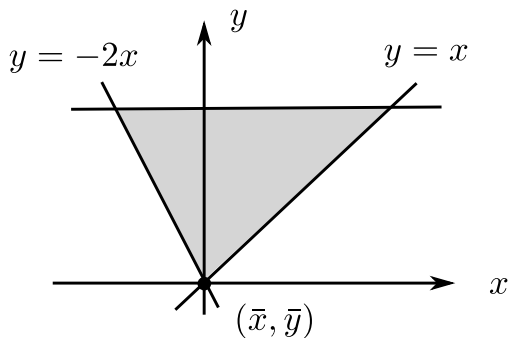
Consider the problem

$$\inf_{(x,y) \in \mathbb{R}^2} f(x,y) := y, \quad \text{s.t.} \quad \begin{cases} g_1(x,y) := 2x + y \geq 0, \\ g_2(x,y) := -x + y \geq 0, \\ g_3(x,y) := -y + 3 \geq 0. \end{cases}$$

- Draw the feasible set and find (geometrically) the solution.
- Verify that the KKT conditions are satisfied.

Linear constraints

Solution.



Linear constraints

- Solution to the problem: $(\bar{x}, \bar{y}) = 0$.
- Let $\lambda \in \mathbb{R}^3$ be the associated Lagrange multiplier.
Necessarily $\lambda_3 = 0$, since $g_3(\bar{x}, \bar{y}) > 0$, by complementarity.
- Lagrangian:

$$L(x, y, \lambda) = y - \lambda_1(2x + y) - \lambda_2(-x + y).$$

- The stationarity condition yields:

$$0 = \frac{\partial L}{\partial x}(0, 0) = -2\lambda_1 + \lambda_2$$

$$0 = \frac{\partial L}{\partial y}(0, 0) = 1 - \lambda_1 - \lambda_2.$$

- This linear system has a unique solution

$$\lambda_1 = 1/3 \geq 0 \quad \lambda_2 = 2/3 \geq 0.$$

The sign condition is satisfied.

Linear constraints

Example 1. Case of **one equality constraint**:

$$m = 1, \quad \mathcal{E} = \{1\}, \quad \mathcal{I} = \emptyset.$$

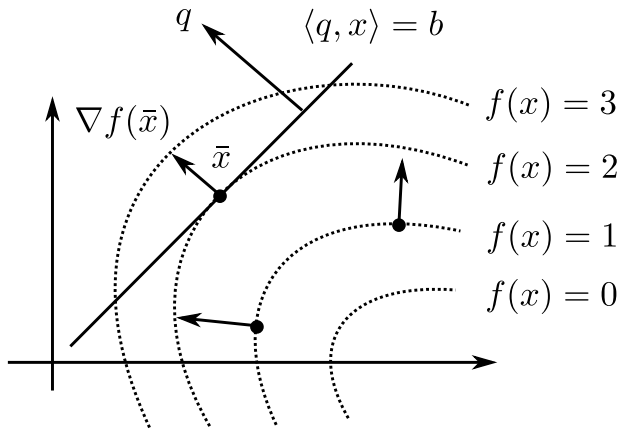
The matrix A is a row vector, let $q = A^\top$.

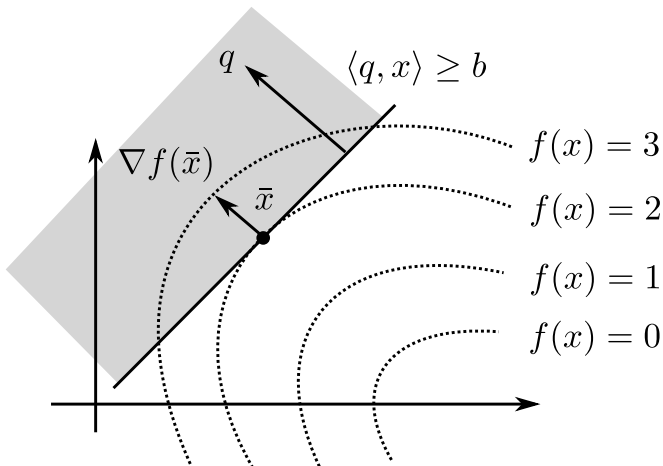
Proof of KKT conditions.

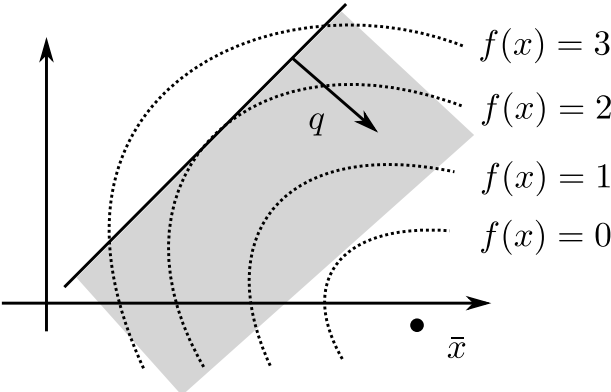
- Geometrically, we understand that $\nabla f(x)$ and q are **colinear**.
- Let $\lambda \in \mathbb{R}$ be such that $\nabla f(x) = \lambda q$.
- We have:

$$\nabla L(x, \lambda) = \nabla f(x) - \lambda \nabla g(x) = \nabla f(x) - \lambda q = 0.$$

Linear constraints







Proof.

- For all $\theta \in [0, \varepsilon/\|h\|]$, we have $\|x_\theta - \bar{x}\| \leq \varepsilon$ and thus

$$f(x_\theta) \geq f(\bar{x}).$$

Linear constraints

- We deduce that

$$0 \leq \lim_{\theta \downarrow 0} \frac{f(\bar{x} + \theta h) - f(\bar{x})}{\theta} = \langle \nabla f(\bar{x}), h \rangle.$$

- Since $h \in \text{Ker}(A)$, we also have $-h \in \text{Ker}(A)$. Therefore,

$$0 \leq \langle \nabla f(\bar{x}), -h \rangle$$

and therefore $\langle \nabla f(\bar{x}), h \rangle = 0$.

- We deduce that

$$\nabla f(x) \in (\text{Ker}(A))^{\perp} = \text{Im}(A^{\top}),$$

that is, there exists $\lambda \in \mathbb{R}^m$ such that $\nabla f(x) = A^{\top} \lambda$.

- We have $\nabla_x L(x, \lambda) = \nabla f(x) - A^{\top} \lambda = 0$.

1 General introduction

- What is an optimization problem?
- Classes of problems
- Existence of a solution
- Derivatives

2 Methods for unconstrained optimization

- Optimality conditions
- Gradient methods
- Newton's method

3 Optimality conditions for constrained problems

- Linear constraints
- **Non-linear constraints**
- Sensitivity analysis

Non-linear constraints

Definition 18

Let \bar{x} be a feasible point. Let the set of **active inequality constraints** $\mathcal{I}_0(\bar{x})$ be defined by

$$\mathcal{I}_0(\bar{x}) = \{i \in \mathcal{I} \mid g_i(\bar{x}) = 0\}.$$

We say that the **Linear Independence Qualification Condition (LICQ)** holds at \bar{x} , if the following set of vectors is linearly independent:

$$\{\nabla g_i(\bar{x})\}_{i \in \mathcal{E} \cup \mathcal{I}_0(\bar{x})}.$$

Theorem 19

Let \bar{x} be a local solution to (P). Assume that the LICQ holds at \bar{x} . Then there exists a unique λ such that the KKT conditions are satisfied.

Remarks.

- Many available **variants** of this theorem in the literature, with different qualification conditions.
- At a numerical level, a solution that does not satisfy the LICQ is hard to compute.

Non-linear constraints

Example 4.

Consider the problem

$$\inf_{x \in \mathbb{R}} x, \quad \text{subject to: } -x^2 \geq 0.$$

Unique feasible point: $\bar{x} = 0$, thus the solution.

Lagrangian:

$$L(x, \lambda) = x + \lambda x^2.$$

At zero:

$$\nabla_x L(0, \lambda) = 1 + 2\lambda \bar{x} = 1 \neq 0.$$

The LICQ is not satisfied, since $\nabla g_1(0) = 0$.

Non-linear constraints

Theorem 20

Assume that

- f is convex
- for all $i \in \mathcal{E}$, the map $x \mapsto g_i(x)$ is affine
- for all $i \in \mathcal{I}$, the map $x \mapsto -g_i(x)$ is convex.

Then any feasible point \bar{x} satisfying the KKT conditions is a global solution to the problem.

Remark. The result holds whether the LICQ holds or not at \bar{x} .

-

- _____

Conclusion

1. *Journal of the American Medical Association*, 2000; 284: 1039-1044.

$$\begin{aligned} f(x, y) &\geq x + y^2 + y + x^2 \\ &= \frac{1}{2}(x^2 + y^2) + \frac{1}{2}(x^2 + 2x + 1) + \frac{1}{2}(y^2 + 2y + 1) - 1 \\ &= \frac{1}{2}\|(x, y)\|^2 + (x + 1)^2 + (y + 1)^2 - 1 \xrightarrow{\|(x, y)\| \rightarrow \infty} \infty. \end{aligned}$$

Exercise

2. It holds:

$$\frac{\partial f}{\partial x} = \exp(x + y^2) + 2x, \quad \frac{\partial f}{\partial y} = 2y \exp(x + y^2) + 1.$$

Therefore, $\nabla f(x, y) = \begin{pmatrix} \exp(x + y^2) + 2x \\ 2y \exp(x + y^2) + 1 \end{pmatrix}.$

We also have

$$\frac{\partial^2 f}{\partial x^2} = \exp(x + y^2) + 2, \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 2y \exp(x + y^2),$$

$$\frac{\partial^2 f}{\partial y^2} = 2 \exp(x + y^2) + 4y^2 \exp(x + y^2).$$

Thus, $D^2 f(x, y) = \begin{pmatrix} \exp(x + y^2) + 2 & 2y \exp(x + y^2) \\ 2y \exp(x + y^2) & (2 + 4y^2) \exp(x + y^2) \end{pmatrix}.$

3. Proof of positive definiteness of D^2f . It holds:

$$a + c = (3 + 4y^2) \exp(x + y^2) + 2 \geq 0$$

and

$$ac - b^2 = 2 \exp(2x + 2y^2) + 4(1 + 2y^2) \exp(x + y^2) \geq 0.$$

It follows that $D^2f(x, y)$ is positive semidefinite, for all (x, y) . Therefore f is a convex function.

Exercise

4. Feasibility of $(0, 0)$: we have $0 + 0 \geq 0$ and $0 + 2 > 0$.

KKT conditions. Lagrangian:

$$L(x, y, \lambda_1, \lambda_2) = \exp(x + y^2) + y + x^2 - \lambda_1(x + y) - \lambda_2(x + 2).$$

Therefore,

$$\frac{\partial L}{\partial x}(0, 0, \lambda_1, \lambda_2) = 1 - \lambda_1 - \lambda_2, \quad \frac{\partial L}{\partial y}(0, 0) = 1 - \lambda_1.$$

Taking $\lambda_1 = 1$ and $\lambda_2 = 0$, we have:

- 1 Stationarity: $\frac{\partial L}{\partial x}(0, 0, 1, 0) = \frac{\partial L}{\partial y}(0, 0, 1, 0) = 0$.
- 2 Sign condition: $\lambda_1 \geq 0$, $\lambda_2 \geq 0$.
- 3 Complementarity: the second constraint is inactive and the corresponding Lagrange multiplier is null.

5. We have the following:
- The cost function is convex.
 - The functions $-(x + y)$ and $-(x + 2)$ are convex.
 - The point $(0, 0)$ is feasible and satisfies the KKT conditions.

Therefore $(0, 0)$ is a global solution.

Exercise:

-

Non-linear constraints

Verification of the LICQ.

$$\nabla g_1(\bar{x}) = \begin{pmatrix} -2\bar{x}_1 \\ -4\bar{x}_2 \end{pmatrix} = \begin{pmatrix} -2 \\ -4 \end{pmatrix} \quad \text{and} \quad \nabla g_2(\bar{x}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

We have: $\mathcal{E} = \emptyset$, $\mathcal{I}_0(\bar{x}) = \{1, 2\}$. The vectors $\nabla g_1(\bar{x})$ and $\nabla g_2(\bar{x})$ are linearly independent, since

$$\det \begin{pmatrix} -2 & -4 \\ -1 & 0 \end{pmatrix} = -4 \neq 0.$$

Thus the LICQ is satisfied at \bar{x} .

Non-linear constraints

KKT conditions.

- Lagrangian:

$$L(x, \lambda) = (-x_1 - x_2) - \lambda_1(-x_1^2 - 2x_2^2 + 3) - \lambda_2(-x_1 + 1).$$

- Stationarity condition:

$$\begin{pmatrix} -1 \\ -1 \end{pmatrix} + \begin{pmatrix} 2\bar{x}_1 \\ 4\bar{x}_2 \end{pmatrix} \lambda_1 + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \lambda_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

It is satisfied at \bar{x} with $\lambda_1 = 1/4 \geq 0$ and $\lambda_2 = 1/2 \geq 0$.

- The sign condition is satisfied.
- The complementarity condition is satisfied (all inequality constraints are active).

1 General introduction

- What is an optimization problem?
- Classes of problems
- Existence of a solution
- Derivatives

2 Methods for unconstrained optimization

- Optimality conditions
- Gradient methods
- Newton's method

3 Optimality conditions for constrained problems

- Linear constraints
- Non-linear constraints
- **Sensitivity analysis**

Sensitivity analysis

- Consider the family of optimization problems

$$\inf_{x \in \mathbb{R}^n} f(x), \quad \text{s.t.} \quad \begin{cases} g_i(x) = y_i, & \forall i \in \mathcal{E}, \\ g_i(x) \geq y_i, & \forall i \in \mathcal{I}, \end{cases} \quad (P(y))$$

parametrized by the vector $y \in \mathbb{R}^m$.

- Let the **value function** V be defined by

$$V(y) = \text{val}(P(y)).$$

Sensitivity analysis

Theorem 21

Assume that for some \bar{y} , the problem $(P(\bar{y}))$ has a solution \bar{x} satisfying the KKT conditions. Let λ denote the corresponding Lagrange multiplier.

Then, under some technical assumptions, V is differentiable at \bar{y} and

$$\nabla V(\bar{y}) = \lambda.$$

Interpretation. A variation δy_i in the i -th constraint generates a variation of the optimal cost of $\lambda_i \delta y_i$ (in **first approximation**).

Sensitivity analysis

Exercise.

A company decides to rent an engine over d days. The engine can be used to produce two different objects. The two objects are not produced simultaneously. Let x_1 and x_2 denote the times dedicated to the production of each object. The resulting benefits (in k€) are given by:

$$\frac{x_1}{1 + x_1} \quad \text{and} \quad \frac{x_2}{4 + x_2}.$$

Sensitivity analysis

- 1 Formulate the problem as a minimization problem.
- 2 Justify the existence of a solution.
- 3 Write the KKT conditions. What is the unit of the dual variable?
- 4 Verify that $\bar{x} = (4, 6)$ satisfies the KKT conditions for $d = 10$ days. Is it a global solution to the problem?
- 5 The renting cost of the engine is 70€/day. Is it of interest for the company to rent the engine for a longer time?

Sensitivity analysis

1. Problem:

$$\inf_{x \in \mathbb{R}^2} -\frac{x_1}{1+x_1} - \frac{x_2}{4+x_2}, \quad \text{s.t.} \quad \begin{cases} x_1 + x_2 = d \\ x_1 \geq 0 \\ x_2 \geq 0 \end{cases}$$

2. The feasible set is obviously compact and non-empty and the cost function is continuous. Therefore, there exists a solution.

Sensitivity analysis

3. Let \bar{x} be a solution. Let $\lambda \in \mathbb{R}^3$ be the associated Lagrange multiplier. Lagrangian:

$$L(x, \lambda) = -\frac{x_1}{1+x_1} - \frac{x_2}{4+x_2} - \lambda_1(x_1 + x_2 - d) - \lambda_2 x_1 - \lambda_3 x_2.$$

KKT conditions:

- Stationarity:

$$-\frac{1}{(1+\bar{x}_1)^2} - \lambda_1 - \lambda_2 = 0, \quad -\frac{4}{(4+\bar{x}_2)^2} - \lambda_1 - \lambda_3 = 0.$$

- Sign condition: $\lambda_2 \geq 0, \lambda_3 \geq 0$.
- Complementarity: $\bar{x}_1 > 0 \Rightarrow \lambda_2 = 0, \bar{x}_2 > 0 \Rightarrow \lambda_3 = 0$.

- Units: $[\lambda_1] = [\lambda_2] = [\lambda_3] = \text{k€}/\text{day}$.

Sensitivity analysis

4. Let λ be such that the KKT conditions hold true. By complementarity condition, we necessarily have $\lambda_2 = \lambda_3 = 0$. The stationarity condition holds true with

$$\lambda_1 = -\frac{1}{(1 + \bar{x}_1)^2} = -\frac{4}{(4 + \bar{x}_2)^2} = -\frac{1}{25} = -0.04.$$

The sign condition trivially holds true since the inequality constraints are inactive.

The point \bar{x} is feasible and satisfies the KKT conditions. We have affine constraints and a convex cost function, therefore, the KKT conditions are sufficient. The point \bar{x} is a global solution.

Sensitivity analysis

5. Increasing the renting time of y days will generate a variation of cost of $\lambda_1 y$ (approximately), that is, an augmentation of the benefit of 40€/day (less the renting price). It would be of interest for the company to reduce the renting time.