

Continuous optimization

ENT305A

Laurent Pfeiffer

Inria and CentraleSupélec, Université Paris-Saclay

Ensta-Paris

Institut Polytechnique de Paris

December 2023

Inria



CentraleSupélec

université
PARIS-SACLAY

© 2010 Pearson Education, Inc. or its affiliate(s). All rights reserved.

0

- 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).

Discussion

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.

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.

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

-

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

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

- 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:

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

[illegible]

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

—

1. *Chlorophyll a* and *Chlorophyll b* were determined by the method of Arar and Collins (1971).

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

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

CONCLUSIONS

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

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*, 1990, 27, 1, 1-14.

- 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.

1 2 3 4 5

- Mechanical structures
- Electricity networks
- Gas networks

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

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

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

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

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

1. *Journal of the American Medical Association*, 1997; 277: 1039-1043.

(2)

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

- K is non-empty and compact (i.e. closed and bounded)
- f is continuous on K .

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. In practical exercises, it is not necessary to justify the continuity of g .

[illegible]

Exercise. Consider

$$f: (x, y) \in \mathbb{R}^2 \mapsto x^4 - 2xy + 2y^2.$$

Prove that f is coercive on \mathbb{R}^2 .

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

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.

Remark. This is a generalization of the previous lemma, since if K is bounded, f is necessarily coercive.

(1) \mathcal{C}_1 is a \mathbb{C}^* -extension of \mathcal{C}_0 if and only if \mathcal{C}_1 is a \mathbb{C}^* -extension of \mathcal{C}_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 7

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{C}_1 is a \mathbb{C}^* -extension of \mathcal{C}_0 if and only if \mathcal{C}_1 is a \mathbb{C}^* -extension of \mathcal{C}_0 .

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

() () () ()

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

Derivatives

Definition 9

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.

Calculate the gradient and the Hessian of the function

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 the American Medical Association*, 1997; 277: 1001-1005.

Optimality conditions

Theorem 11

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 12

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 13

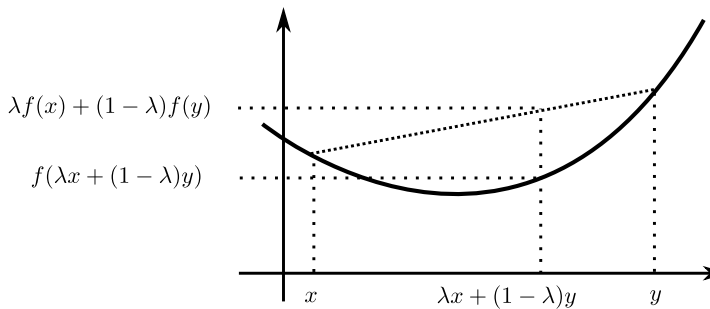
- *The function f is convex if and only if*

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

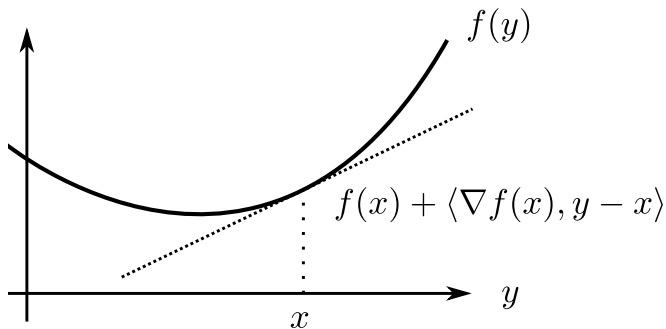
for all x and $y \in \mathbb{R}^n$.

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

Optimality conditions



Optimality conditions



1. *Journal of the American Medical Association*, 1997; 278: 1039-1044.

1. *Journal of Management Studies*, 1996, 33(1), 1-14.

Journal Pre-proof

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ≡ ↺ 🔍 ↻

Journal Pre-proof

$$x \text{ is stationary} \iff \nabla f(x) = 0$$

$$\iff Ax + b = 0$$

$$\Leftrightarrow x = -A^{-1}b.$$

Therefore there is a unique stationary point, which concludes the proof.

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

© 2000 Blackwell Science 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 15

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,

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

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

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

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

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^n \quad (x, y) \mapsto (x + y, x - y)$$

[illegible]

1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 2678, 2679, 2680, 26

$$\phi'_k(\alpha) = \langle \nabla f(x_k + \alpha d_k), d_k \rangle.$$

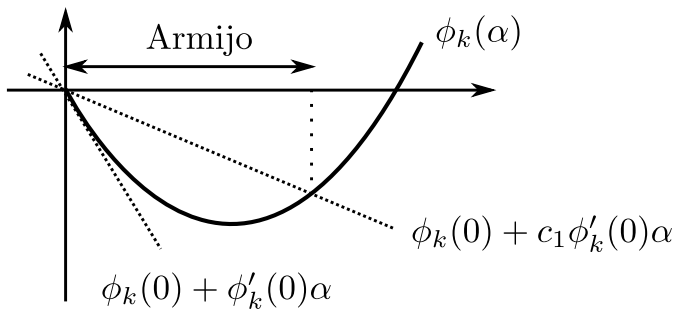
In particular, since d_k is a descent direction,

$$\phi'_k(0) = \langle \nabla f(x_k), d_k \rangle < 0.$$

Definition 16

Let us fix $0 < c_1 < 1$. We say that α satisfies **Armijo's rule** if

$$\phi_k(\alpha) \leq \phi_k(0) + c_1 \phi'_k(0) \alpha.$$

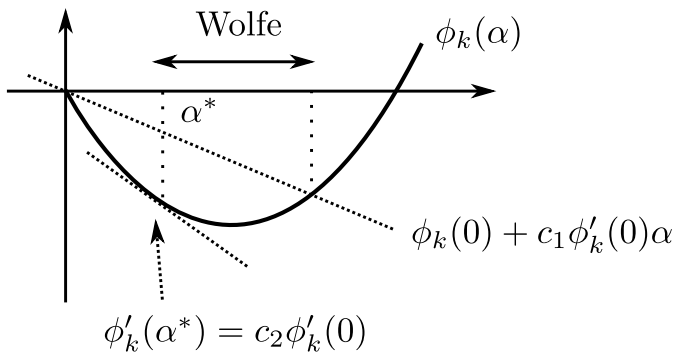


Definition 17

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



Discussion

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

100

Given α , we have

Journal of Management Inquiry 18(6)

(1) $\frac{1}{2} \leq \frac{a}{b} \leq \frac{3}{2}$

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:

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

The difficulties mentioned above are still relevant.

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

where

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

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

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

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

- Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ and let $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ be two continuously differentiable functions.
- Let $(\mathcal{E}, \mathcal{I})$ be a partition of $\{1, \dots, m\}$ (that is, $\mathcal{E} \cap \mathcal{I} = \emptyset$, $\mathcal{E} \cup \mathcal{I} = \{1, \dots, m\}$).
- Let the **Lagrangian** $L: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ be defined by

$$L(x, \lambda) = f(x) - \langle \lambda, g(x) \rangle = f(x) - \sum_{i=1}^m \lambda_i g_i(x).$$

The variable λ is referred to as **dual variable**.

Linear equality constraints

- We investigate in this section the problem

$$\inf_{x \in \mathbb{R}^n} f(x), \quad \text{s.t.} \quad \begin{cases} g_i(x) = 0, & \forall i \in \mathcal{E} \\ g_i(x) \geq 0, & \forall i \in \mathcal{I}. \end{cases} \quad (P)$$

- Let $x \in \mathbb{R}^n$ be feasible. Let $i \in \mathcal{I}$. We say that
 - the inequality constraint i is **active** if $g_i(x) = 0$
 - the inequality constraint i is **inactive** if $g_i(x) > 0$.
- *Remark.* All results of the section are true if $\mathcal{E} = \emptyset$ or $\mathcal{I} = \emptyset$.

Linear equality constraints

Theorem 18

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 $i \in \mathcal{I}$,
 $g_i(\bar{x}) > 0 \implies \lambda_i = 0$.

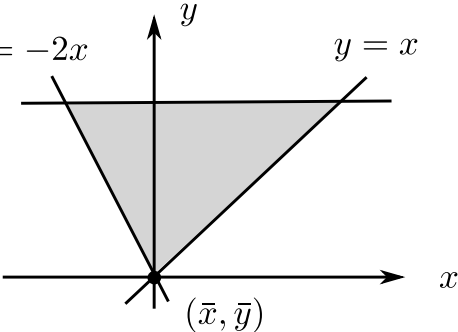
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840. 84

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.



[illegible]

24

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

The sign condition is satisfied.

Linear constraints

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

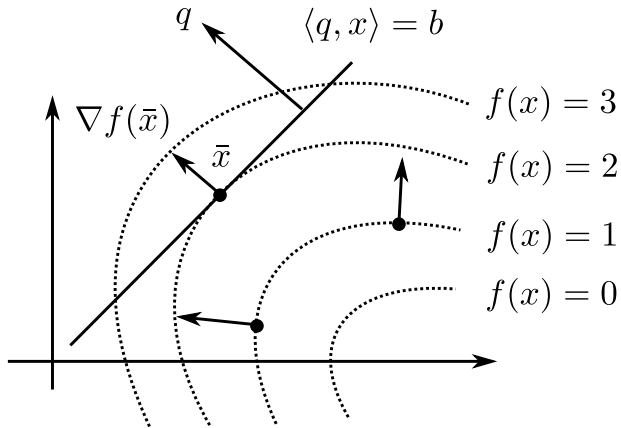
$$m = 1, \quad \mathcal{E} = \{1\}, \quad \mathcal{I} = \emptyset.$$

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

Proof of KKT conditions.

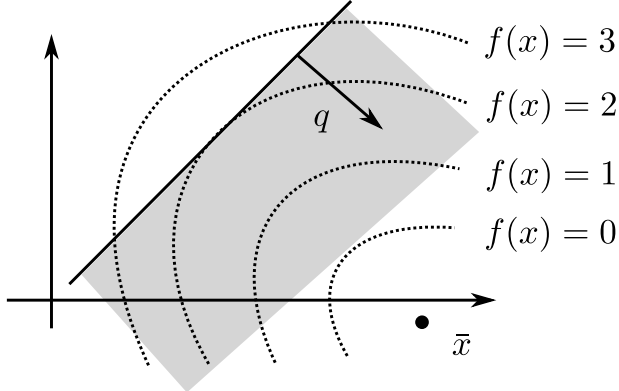
- Geometrically, we understand that $\nabla f(x)$ and q are **colinear**.
- Let $\lambda \in \mathbb{R}$ be such that $\nabla f(x) = \lambda q$.
- We have:

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



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





Linear constraints

Example 3. Case of m equality constraints ($\mathcal{I} = \emptyset$).

Proof.

- Let $\varepsilon > 0$ be given by the definition of a local solution.
Let $h \in \text{Ker}(A)$ (that is $Ah = 0$).
For all $\theta \in \mathbb{R}$, let $x_\theta = \bar{x} + \theta h$.
- For all $\theta \in \mathbb{R}$, x_θ is **feasible**:

$$g(x_\theta) = Ax_\theta + b = A\bar{x} + b + \theta Ah = 0.$$

- 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

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

$$\left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 0 & 1 \end{pmatrix} \right)$$

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

© 2010 Pearson Education, Inc. or its affiliate(s). All rights reserved.

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 $\nabla g_1(0) = 0$.

- 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} .

Exercise

Exercise. Consider the function $f: (x, y) \in \mathbb{R}^2 \mapsto \exp(x + y^2) + y + x^2$.

- 1 Prove that f is coercive.
- 2 Calcule $\nabla f(x, y)$ and $\nabla^2 f(x, y)$.
- 3 We recall that a symmetric matrix of size 2 of the form $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ is positive semidefinite if and only if $a + c \geq 0$ and $ac - b^2 \geq 0$. Using this fact, prove that f is convex.
- 4 We consider the following problem:

$$\inf_{(x,y) \in \mathbb{R}^2} f(x, y), \quad \text{subject to: } \begin{cases} x + y \geq 0 \\ x + 2 \geq 0. \end{cases} \quad (\mathcal{P})$$

Verify that $(0, 0)$ is feasible and satisfies the KKT conditions.

- 5 Is the point $(0, 0)$ a global solution to problem (\mathcal{P}) ?

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

()

11/11/2019 11:11 AM

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.

Journal of Management Inquiry 18(6)

Introduction

Conclusions

-

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 22

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.