RAMP-UP "DATA SCIENCE"
PART: CONTINUOUS OPTIMIZATION

CHRISTIAN KIRCHES

1. Constrained Optimization: Quick Questions

(1) Define 'feasible set' and 'feasible point'
(2) Define 'active inequality' and 'active set'
(3) How many active sets are there if you're given $m \geqslant 1$ inequalities?
(4) Define 'linearized cone for the feasible set at a point'
(5) Define 'global', 'local', and 'strict local minimum' of a constrained nonlinear optimization problem
(6) Define 'lagrangian function'
(7) Define 'lagrange multiplier'
(8) Give a necessary optimality conditions of first order for constrained optimization

2. Using the Lagrange (KKT) Formalism A

Determine all KKT points (x^*, λ^*) of the minimization problem for the objective $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = -x_1^3 x_2^5$ subject to the equality constraint $g(x) = x_1 + x_2 - 8 = 0$.

3. Using the Lagrange (KKT) Formalism B

Consider the unit sphere in $\mathbb{R}^3$, which is named $\mathbb{S}^2$ (because it has a two-dimensional surface) and is given by the set

$$\mathbb{S}^2 := \left\{ (x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1 \right\}.$$

Use the Lagrange function and formalism to determine the largest and smallest (Euclidean) distance between the point $b = (2, 1, -2)^T \in \mathbb{R}^3$ and the sphere $\mathbb{S}^2$.

4. Active Sets

Consider the constrained nonlinear minimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & \tfrac{1}{2}(x_1 - 2)^2 + \tfrac{1}{2}(x_2 - \tfrac{1}{2})^2 \\ \text{s.t.} \quad & (x_1 + 1)^{-1} - x_2 - \tfrac{1}{4} \geqslant 0, \\ & x_1 \geqslant 0, \\ & x_2 \geqslant 0. \end{aligned}$$

Which and how many possibilities are there for picking active sets $\mathcal{A}(x)$?

Find a local minimum by trying to verify the KKT conditions for all active sets, one by one.

5. Extrema of Inequality-Constrained Nonlinear Problems A

Consider the constrained nonlinear minimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & f(x) = (x_1 - 2)^2 + (x_2 - 1)^2 \\ \text{s.t.} \quad & x_2 \leqslant 2 - x_1 \\ & x_2 \geqslant x_1^2 \end{aligned}$$

Determine the functions $h_i(x) \geqslant 0$ that represent the inequality constraints.

Date: June 20, 2023.

Draw a sketch of the feasible set of this minimization problem. Also draw some level sets of the objective function f.

In the first quadrant ($x_1 \geqslant 0$, $x_2 \geqslant 0$), find the location $\bar{x}$ with $h_1(\bar{x}) = h_2(\bar{x})$ and indicate the gradients $\nabla h_1(\bar{x})$, $\nabla h_2(\bar{x})$, and the antigradient $-\nabla f(\bar{x})$ by arrows.

By reasoning about your sketch, determine all KKT points and solve the minimization problem. You don't have to carry out symbolic calculations.

6. Extrema of Inequality-Constrained Nonlinear Problems B

Let $u > 0$ be the circumference of a triangle with sides of lengths $x, y, z \geqslant 0$. To obtain a proper triangle, we must have $x, y, z \leqslant \frac{u}{2}$.

Determine the particular triangle that has maximum area. Use Heron's formula

$$F(x, y, z) = \sqrt{\frac{u}{2}\left(\frac{u}{2} - x\right)\left(\frac{u}{2} - y\right)\left(\frac{u}{2} - z\right)}$$

for the area.

First come up with a suitable objective function, then add the constraint you require.

Finally, form the Lagrangian function and determine all KKT points to find the maximum area and the associated configuration of the triangle.

Solutions

Quick Questions

(1) For the NLP $\min\ f(x)$ s.t. $g(x) = 0,\ h(x) \geqslant 0$, the feasible set is defined as $\mathcal{F} := \{x \in \mathbb{R}^n \mid g(x) = 0,\ h(x) \geqslant 0\}$. A point $\bar{x} \in \mathbb{R}^n$ is called a feasible point if $\bar{x} \in \mathcal{F}$, equivalently if $g(\bar{x}) = 0$ and $h(\bar{x}) \geqslant 0$.

(2) An inequality $h_i(x) \geqslant 0$ is called active in a point $\bar{x} \in \mathbb{R}^n$ if $h_i(\bar{x}) = 0$ holds. The active set at $\bar{x}$ is the set of indices $1 \leqslant i \leqslant m$ of inequalities active in $\bar{x}$, i.e., $\mathcal{A}(\bar{x}) = \{1 \leqslant i \leqslant m \mid h_i(\bar{x}) = 0\}$.

(3) There are 2^k active sets for k inequalities.

(4) The linearized cone of the feasible set $\mathcal{F}$ at $\bar{x} \in \mathcal{F}$ is defined as $\mathcal{L}(\mathcal{F}, \bar{x}) := \{d \in \mathbb{R}^n \mid \nabla g(x)^T d = 0,\ \nabla h_i(x)^T d \geqslant 0,\ i \in \mathcal{A}(\bar{x})\}$.

(5) See the essentials' sheet.

(6) The Lagrangian function associated with the NLP $\min\ f(x)$ s.t. $g(x) = 0,\ h(x) \geqslant 0$ is defined as $\mathcal{L}(x, \lambda, \mu) = f(x) - \lambda^T g(x) - \mu^T h(x)$.

(7) The vectors $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^k$ in the Lagrangian function are called Lagrange multipliers.

(8) There are multiple valid answers:
   (a) Let $\bar{x}$ be a local minimum of (NLP) [and let a CQ hold in $\bar{x}$]. Then the following holds: $\nabla f(\bar{x})^T d \geqslant 0$ for all $d \in \mathcal{L}(x, \lambda, \mu)$.
   (b) Let $\bar{x}$ be a local minimum of (NLP) [and let a CQ hold in $\bar{x}$]. Then the following holds: $\nabla f(\bar{x})^T d \leqslant 0$ for all $d \in \mathcal{L}(x, \lambda, \mu)^\circ$.
   (c) Let $\bar{x}$ be a local minimum of (NLP) [and let a CQ hold in $\bar{x}$]. Then there exists vectors $\bar{\lambda} \in \mathbb{R}^m$ and $\bar{\mu} \in \mathbb{R}^k$ such that the following holds: $\nabla_x \mathcal{L}(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$ (stationarity), $g(\bar{x}) = 0$ and $h(\bar{x}) \geqslant 0$ (feasibility), $\bar{\mu} \geqslant 0$ (optimality/dual feasibility), $\bar{\mu}^T h(\bar{x}) = 0$ (complementary slackness).

Using the Lagrange (KKT) Formalism A

$$\begin{aligned}
\mathcal{L}(x, \lambda) &= -x_1^3 x_2^5 - \lambda(x_1 + x_2 - 8) \\
\nabla_x \mathcal{L}(x, \lambda) &= \begin{pmatrix} -3x_1^2 x_2^5 - \lambda \\ -5x_1^3 x_2^4 - \lambda \end{pmatrix}
\end{aligned}$$

Solving for λ and asking for stationarity yields the polynomial root finding problem

$$\begin{aligned}
0 &= -3x_1^2 x_2^5 + 5x_1^3 x_2^4 \\
&= x_1^2 x_2^4 (5x_1 - 3x_2)
\end{aligned}$$

Three solutions exist: $x_1 = 0$, x_2 arbitrary; $x_2 = 0$, x_1 arbitrary; $x_1 = \frac{3}{5} x_2$. Asking for feasibility narrows this down to the KKT points

$$x^{(1)} = (0, 8)^T, \quad x^{(2)} = (8, 0)^T, \quad x^{(3)} = (3, 5)^T.$$

Using the Lagrange (KKT) Formalism B

The objective to minimize/maximize is the Euclidean distance $||(x_1, x_2, x_3)^T - (2, 1, -2)^T||_2$. This quantity can be squared without moving the location of the minimizer, hence

$$f(x) = (x_1 - 2)^2 + (x_2 - 1)^2 + (x_3 + 2)^2.$$

This leads to

$$\begin{aligned}
\mathcal{L}(x, \lambda) &= (x_1 - 2)^2 + (x_2 - 1)^2 + (x_3 + 2)^2 - \lambda(x_1^2 + x_2^2 + x_3^2 - 1) \\
\nabla_x \mathcal{L}(x, \lambda) &= \begin{pmatrix} 2(x_1 - 2) - 2\lambda x_1 \\ 2(x_2 - 1) - 2\lambda x_2 \\ 2(x_3 + 2) - 2\lambda x_3 \end{pmatrix}
\end{aligned}$$

Solving for the x_i yields identities in terms of λ:

$$\begin{aligned} x_1 &= 2/(1-\lambda) \\ x_2 &= 1/(1-\lambda) \\ x_3 &= -2/(1-\lambda) \end{aligned}$$

Inserting into the constraint to ensure feasibility yields

$$0 = x_1^2 + x_2^2 + x_3^2 - 1 = \frac{9}{(1-\lambda)^2} - 1$$

with the only solutions being $\lambda \in \{-2, +4\}$. Inserting into the identities for the x_i yields the two KKT points

$$x^* = \left(\frac{2}{3}, \frac{1}{3}, -\frac{2}{3}\right)^T \text{ and } \hat{x} = \left(-\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}\right)^T.$$

The maximum distance is obtained in $\hat{x}$, the minimum one in x^*.

Active Sets

$$\begin{aligned} \mathcal{L}(x,\mu) &= \tfrac{1}{2}(x_1-2)^2 + \tfrac{1}{2}(x_2-\tfrac{1}{2})^2 - \mu_1((x_1+1)^{-1} - x_2 - \tfrac{1}{4}) - \mu_2 x_1 - \mu_3 x_2 \\ \nabla_x \mathcal{L}(x,\mu) &= \begin{pmatrix} x_1 - 2 + \mu_1((x_1+1)^{-2}) - \mu_2 \\ (x_2 - \tfrac{1}{2}) + \mu_1 - \mu_3 \end{pmatrix} \end{aligned}$$

The unconstrained minimizer would be $\bar{x} = (2, \frac{1}{2})^T$, but is infeasible with respect to the first inequality,

$$(\bar{x}_1+1)^{-1} - \bar{x}_2 - \tfrac{1}{4} = (2+1)^{-1} - \tfrac{1}{2} - \tfrac{1}{4} = \frac{4-6-3}{12} < 0.$$

From this, we draw the conclusion that the first inequality is active,

$$x_2 = (x_1+1)^{-1} - \tfrac{1}{4}.$$

Assumption $x_1 = 0$ and $x_2 = 0$.
This active set is infeasible for the first inequality constraint, which we assumed to be active.

Assumption $x_1 = 0$.
Then $x_2 = (0+1)^{-1} - \frac{1}{4} = \frac{3}{4} > 0$, such that $\mu_3 = 0$ due to complementarity. The second row of stationarity yields

$$(\tfrac{3}{4} - \tfrac{1}{2}) + \mu_1 = 0$$

such that $\mu_1 = -\frac{1}{4} < 0$, which violates the optimality/dual feasibility KKT condition.

Assumption $x_2 = 0$.
Then $0 = (x_1+1)^{-1} - \frac{1}{4}$, which yields $x_1 = 3$ such that $\mu_2 = 0$ due to complementarity. The second row of stationarity yields

$$(0 - \tfrac{1}{2}) + \mu_1 = 0$$

such that $\mu_1 = \frac{1}{2}$.
Verifying the first row of stationarity fails:

$$0 = 3 - 2 + \tfrac{1}{2}((3+1)^{-2}) - 0 = 1\tfrac{1}{32}.$$

Vice versa, we could have computed a negative μ_1 from the first row of stationarity.

Assumption $x_1 > 0$ and $x_2 > 0$.
Then $\mu_2 = 0$ and $\mu_3 = 0$. Inserting x_2 into the second row of the stationarity condition yields (this is where the sign error was)

$$(\tfrac{3}{4} - (x_1+1)^{-1}) = \mu_1.$$

The first row of stationarity now yields

$$x_1 + (\tfrac{3}{4} - (x_1+1)^{-1})((x_1+1)^{-2}) = 2.$$

This may be transformed into a fourth order polynomial with analytical roots,

$$x_1^4 + x_1^3 - 3x_1^2 - \tfrac{17}{4}x_1 - \tfrac{9}{4}.$$

It has two real soutions $x_1 \approx -1.7467$ (this is infeasible for the second inequality) and $x_1 \approx 1.9528$, which we can only find using a symbolic calculator or a computer. Inserting the remaining feasible solution to find μ_1 yields a positive multiplier,

$$\mu_1 = \tfrac{3}{4} - (1.9528 + 1)^{-1} \approx 0.4113 > 0.$$

Now for x_2 we approximately find

$$x_2 = (1.9528 + 1)^{-1} - \tfrac{1}{4} \approx 0.0886 > 0$$

in agreement with our assumption $x_2 > 0$.

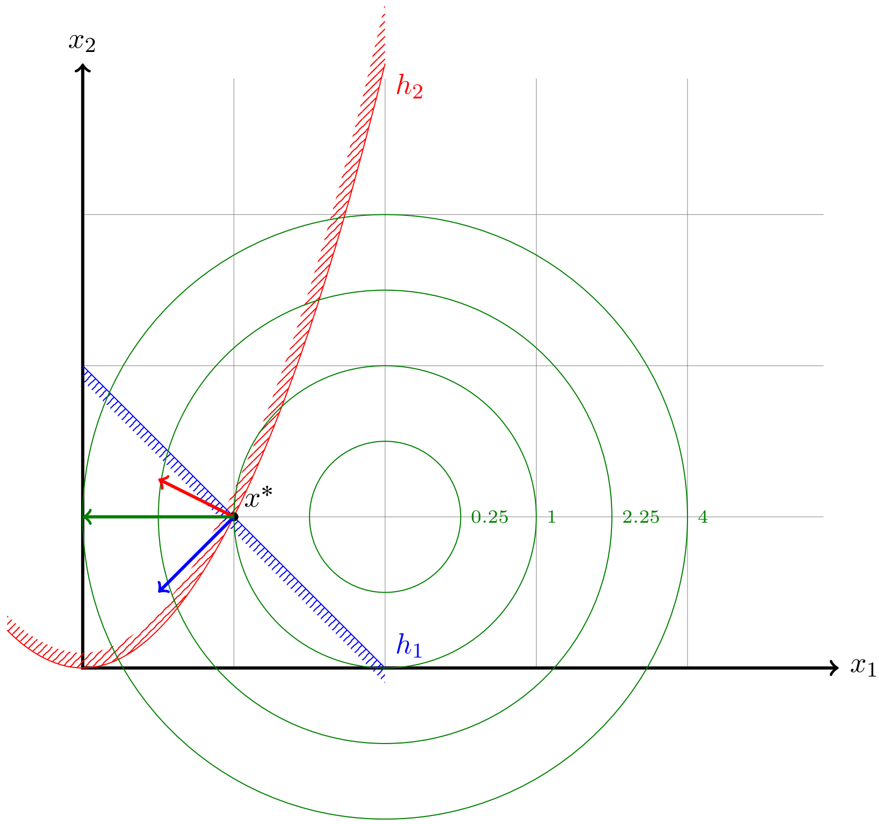
Summarizing, the active set $\mathcal{A} = \{1\}$ admits a KKT point at approximately $x^* = (1.9528, 0.0886)^T$ with approximate multiplier $\mu = (0.4113, 0, 0)^T$.

Extrema of Inequality-Constrained Nonlinear Problems A

$$h_1(x) = 2 - x_1 - x_2 \geqslant 0$$
$$h_2(x) = x_2 - x_1^2 \geqslant 0$$

$$\begin{aligned} & h_1(\bar{x}) = h_2(\bar{x}) \\ \Leftrightarrow \quad & \bar{x}_1(x_1 - 1) = 2(\bar{x}_2 - 1) \end{aligned}$$

The solution in the upper right quadrant is $x^* = (1, 1)^T$.



We immediately see that a linear combination of the antigradient using the active constraints' normals in x^* is possible and uses positive linear factors (Lagrange multipliers μ). Thus x^* is stationary and optimality/dual feasibility is satisfied. x^* is also feasible. All inequalities are active, so complementary slackness is satisfied. Hence x^* is a KKT point. As the feasible set is convex and the objective is strictly convex, x^* is the unique local and global minimum.

Extrema of Inequality-Constrained Nonlinear Problems B

Square the objective as done above to find the problem

$$\begin{aligned}\min_{x,y,z,u} &\frac{u}{2}\left(\frac{u}{2}-x\right)\left(\frac{u}{2}-y\right)\left(\frac{u}{2}-z\right)\\ \text{s.t.} &x+y+z-u=0\end{aligned}$$

We ignore nonnegativity of x, y, z (this could be taken care of when selecting among the KKT points we found) and properness of the triangle (an improper triangle would have zero area). We only compute the Lagrange gradient w.r.t. (x, y, z) to save some work:

$$\mathcal{L}(x,y,z,\lambda) = \frac{u}{2}\left(\frac{u}{2}-x\right)\left(\frac{u}{2}-y\right)\left(\frac{u}{2}-z\right) - \lambda(x+y+z-u)$$

$$\nabla_{(x,y,z)}\mathcal{L}(x,y,z,\lambda) = \begin{pmatrix} \frac{1}{8}u(u-2y)(u-2z)-\lambda \\ \frac{1}{8}u(u-2x)(u-2z)-\lambda \\ \frac{1}{8}u(u-2x)(u-2y)-\lambda \end{pmatrix}$$

We find the triple identity

$$8\lambda = u(u-2y)(u-2z) = u(u-2x)(u-2z) = u(u-2x)(u-2y),$$

which is solvable only if $x = y = z$. But then $x = y = z = u/3$ for feasibility. The area is

$$\sqrt{\frac{u}{2}\left(\frac{u}{2}-x\right)\left(\frac{u}{2}-y\right)\left(\frac{u}{2}-z\right)} = \sqrt{\frac{u}{2}\left(\frac{u}{6}\right)^3} = \sqrt{\frac{u^4}{2^4\cdot 3^3}} = \frac{1}{12\sqrt{3}}u^2.$$