Recitation 20

Tuesday November 19, 2024

1 Recap

We consider the problem of constrained optimization, where we aim to minimize a function $f_0(x)$ over $x \in \mathbb{R}^n$ subject to m inequality constraints and p equality constraints:

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, \ldots, m, \\ & h_j(x) = 0, \quad j = 1, \ldots, p. \end{aligned}$$

The set of points x that satisfy these constraints is called the *feasible set*, and any x in this set is a *feasible point*.

1.1 Equality Constraints and the Lagrangian

Starting with only equality constraints, we consider minimizing $f(x)$ subject to $h_j(x) = 0$ for $j = 1, \ldots, p$. At an optimal point $x^\star$, we expect the gradient of the objective function $\nabla f(x^\star)$ to lie in the span of the gradients of the constraints $\nabla h_j(x^\star)$. This can be expressed as:

$$\nabla f(x^\star) + \sum_{j=1}^{p} \nu_j \nabla h_j(x^\star) = 0,$$

where ν_j are the Lagrange multipliers. This leads us to define the *Lagrangian* function:

$$L(x, \nu) = f(x) + \sum_{j=1}^{p} \nu_j h_j(x) = f(x) + \vec{h}(x)^T \vec{\nu}.$$

The *Karush-Kuhn-Tucker (KKT) conditions* for this problem are:

$$\begin{aligned} \nabla_x L(x^\star, \nu^\star) &= 0, \\ h_j(x^\star) &= 0, \quad j = 1, \ldots, p. \end{aligned}$$

These conditions are necessary for optimality under certain regularity conditions, such as the Linear Independence Constraint Qualification (LICQ), where the gradients of the active constraints are linearly independent.

1.2 Quadratic Programming Example

Consider the quadratic programming (QP) problem:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2} x^T A x - b^T x \\ \text{subject to} \quad & Cx = d, \end{aligned}$$

where A is a symmetric positive definite (SPD) $n \times n$ matrix, $b \in \mathbb{R}^n$, C is a $p \times n$ matrix with $p < n$, and $d \in \mathbb{R}^p$. The KKT conditions for this problem are:

$$\begin{aligned} Ax^\star - b + C^T \nu^\star = 0, \\ Cx^\star = d. \end{aligned}$$

which can be re-written in matrix form in terms of an $(n+p) \times (n+p)$ matrix:

$$\begin{pmatrix} A & C^T \\ C & 0I \end{pmatrix} \begin{pmatrix} x^\star \\ \nu^\star \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix}.$$

Solving these equations yields the optimal $x^\star$ and the Lagrange multipliers $\nu^\star$.

The first equation $Ax^\star - b + C^T \nu^\star = 0$ has a nice geometrical interpretation, similar to how we derived the minimum-norm solution of $Cx = d$ previously: it means that $\nabla f = Ax^\star - b \in C(C^T) = N(C)^\perp$, i.e. that $\nabla f \perp N(C)$, but since the feasible set (solving $Cx = d$) is parallel to $N(C)$ this is just the condition that ∇f is perpendicular to the feasible set.

1.3 Incorporating Inequality Constraints

When inequality constraints are added, the problem becomes:

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, \ldots, m, \\ & h_j(x) = 0, \quad j = 1, \ldots, p. \end{aligned}$$

Defining the Lagrangian $L(x, \lambda, \nu) = f_0(x) + \vec{f}(x)^T \vec{\lambda} + \vec{h}(x)^T \vec{\nu}$, with new Lagrange multipliers $\vec{\lambda} \in \mathbb{R}^m$ associated with the inequality constraints, the KKT conditions become:

$$\begin{aligned} \nabla_x L(x^\star, \lambda^\star, \nu^\star) = 0, & \\ f_i(x^\star) \leq 0, & \quad i = 1, \ldots, m, \\ h_j(x^\star) = 0, & \quad j = 1, \ldots, p, \\ \lambda_i^\star \geq 0, & \quad i = 1, \ldots, m, \\ \lambda_i^\star f_i(x^\star) = 0, & \quad i = 1, \ldots, m. \end{aligned}$$

Here, $\lambda_i^\star$ are the Lagrange multipliers associated with the inequality constraints. The condition $\lambda_i^\star f_i(x^\star) = 0$ is known as *complementary slackness*, ensuring that either the constraint is active ($f_i(x^\star) = 0$) with $\lambda_i^\star \geq 0$, or the constraint is inactive ($f_i(x^\star) < 0$) with $\lambda_i^\star = 0$. The Lagrange multipliers λ_i for the inequalities must be ≥ 0 because the "downhill" direction ∇f_0 can only point towards the *infeasible* region $f_i > 0$ ("into the wall").

2 Exercises

1. **Optimization on the Unit Circle and Inequality Constraints**

   Consider the optimization problem:

   $$\text{minimize} \quad f(x, y) = 2x + y$$

   $$\text{subject to} \quad h(x, y) = x^2 + y^2 - 1 = 0.$$

   (a) **Graphical Solution** Draw the contours of f and the constraint $h(x, y) = 0$ (the unit circle). Identify the feasible set and approximate the point(s) where f attains its minimum.

   (b) **Substitution Method** Solve the constraint for one variable and substitute into f, then find the minimum by differentiating and setting the derivative to zero.

   (c) **Lagrange Multipliers Method** Use the method of Lagrange multipliers to find the minimum. Set up the equations:

   $$\nabla f(x, y) + \nu \nabla h(x, y) = 0,$$
   $$h(x, y) = 0,$$

   where ν is the Lagrange multiplier.

   (d) **Inequality Constraints** Now, consider the inequality constraints:

   i. $h(x, y) \leq 0$.

   ii. $h(x, y) \geq 0$.

   For each case, determine what happens to the minimum value of $f(x, y)$ and compare to the KKT conditions.

2. Consider the quadratic function $f : \mathbb{R}^3 \to \mathbb{R}$ given by

   $$f(x_1, x_2, x_3) = 4(x_1^2 + x_2^2 + x_3^2) - (x_1 + x_2 + x_3)^2 - x_1 - 2x_2 - 3x_3$$

   (a) Give a matrix Q and a vector r such that $f(x) = x^T Q x + r^T x$.

   (b) Is the function f bounded below? How can you tell?

   (c) Let f_u be the (unconstrained) minimum value of f. What is its value? Is the optimal solution unique?

   (d) We now add the linear constraint $4x_1 + 5x_2 + 6x_3 = 100$, and let f_c be the (constrained) minimum value. What is f_c? Is the (constrained) optimal solution unique?

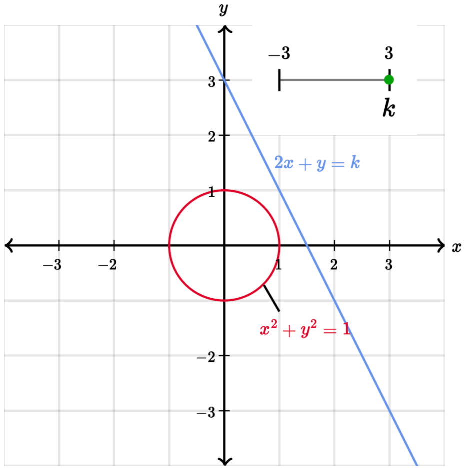
   (e) Give a geometric explanation for why $f_u \leq f_c$.

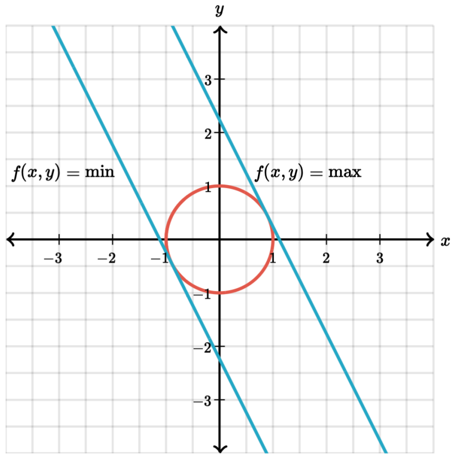
3. If we write the constraints as a vector-valued function $\vec{h}(x) \in \mathbb{R}^p$, the Lagrangian is $L(x, \nu) = f(x) + \vec{h}^T \vec{\nu}$ with a vector $\vec{\nu} \in \mathbb{R}^p$ of Lagrange multipliers. If J is the __ × __ Jacobian matrix of $\vec{h}(x)$, write the $\nabla_x L = \vec{0}$ KKT equation in terms of ∇f, J (or J^T), and $\vec{\nu}$. (A similar procedure works for the inequality constraints.)

3 Solutions

1. **Optimization on the Unit Circle and Inequality Constraints**

   (a) **Graphical Solution** The feasible set is the unit circle $x^2 + y^2 = 1$. The contours of $f(x, y) = 2x + y$ are straight lines. The minimum occurs where the contour line is tangent to the circle in the direction of decreasing f, which is at the point with the smallest possible value of $2x + y$ on the circle.





   (b) **Substitution Method** Solve $y = \pm\sqrt{1 - x^2}$. Substitute into $f(x, y)$:

$$f(x) = 2x + \sqrt{1 - x^2}.$$

   Differentiate $f(x)$ with respect to x:

$$f'(x) = 2 - \frac{x}{\sqrt{1 - x^2}}.$$

   Set $f'(x) = 0$:

$$2 - \frac{x}{\sqrt{1 - x^2}} = 0 \implies \frac{x}{\sqrt{1 - x^2}} = 2.$$

   Solve for x:

$$x^2 = \frac{4}{5} \implies x = \pm\frac{2}{\sqrt{5}}.$$

   Choose $x = -\frac{2}{\sqrt{5}}$ to minimize f. Then $y = -\frac{1}{\sqrt{5}}$. The minimum value is:

$$f_{\min} = 2\left(-\frac{2}{\sqrt{5}}\right) + \left(-\frac{1}{\sqrt{5}}\right) = -\sqrt{5}.$$

   (c) **Lagrange Multipliers Method** Compute the gradients:

$$\nabla f = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \nabla h = \begin{bmatrix} 2x \\ 2y \end{bmatrix}.$$

   Set up the equations:

$$\begin{aligned} 2 + 2x\nu &= 0, \\ 1 + 2y\nu &= 0, \\ x^2 + y^2 - 1 &= 0. \end{aligned}$$

Solve for ν:

$$\nu = -\frac{1}{2x} = -\frac{1}{2y}.$$

Therefore, $2y = x$. Substitute into the constraint:

$$x^2 + \left(\frac{x}{2}\right)^2 = 1 \implies \frac{5x^2}{4} = 1 \implies x = \pm\frac{2}{\sqrt{5}}, y = \pm\frac{1}{\sqrt{5}}$$

Thus, the minimum occurs at $x = -\frac{2}{\sqrt{5}}$, $y = -\frac{1}{\sqrt{5}}$, with $f_{\min} = -\sqrt{5}$.

(d) **Inequality Constraints**

i. $h(x, y) \leq 0$: The feasible set is the interior and boundary of the unit circle. Geometrically, the minimum of f occurs at the same point on the boundry from part (d) $(x, y) = (x^*, y^*) = (-\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}})$, with $f = -\sqrt{5}$. The KKT conditons are

ii. $h(x, y) \geq 0$: The feasible set is the exterior and boundary of the unit circle. As x and y decrease (become more negative), $f(x, y) = 2x + y$ decreases without bound.

Constraint	Lagrangian	KKT Conditions
$h(x) = 0$	$L(x, \nu) = f(x) + \nu h(x)$	$\begin{cases} \nabla f(x) + \nu \nabla h(x) = 0, \\ h(x) = 0. \end{cases} \implies \begin{cases} \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \nu \begin{bmatrix} 2x \\ 2y \end{bmatrix} = 0, \\ x^2 + y^2 - 1 = 0. \end{cases}$
$h(x) \leq 0$	$L(x, \lambda) = f(x) + \lambda h(x)$	$\begin{cases} \nabla f(x) + \lambda \nabla h(x) = 0, \\ h(x) \leq 0, \\ \lambda \geq 0, \\ \lambda h(x) = 0. \end{cases} \implies \begin{cases} \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} 2x \\ 2y \end{bmatrix} = 0, \\ x^2 + y^2 - 1 \leq 0, \\ \lambda \geq 0, \\ \lambda(x^2 + y^2 - 1) = 0. \end{cases}$
$h(x) \geq 0$	$L(x, \lambda) = f(x) - \lambda h(x)$	$\begin{cases} \nabla f(x) - \lambda \nabla h(x) = 0, \\ h(x) \geq 0, \\ \lambda \geq 0, \\ \lambda h(x) = 0. \end{cases} \implies \begin{cases} \begin{bmatrix} 2 \\ 1 \end{bmatrix} - \lambda \begin{bmatrix} 2x \\ 2y \end{bmatrix} = 0, \\ x^2 + y^2 - 1 \geq 0, \\ \lambda \geq 0, \\ \lambda(x^2 + y^2 - 1) = 0. \end{cases}$

2. **Equality Constrained Quadratic Minimization**

(a) The function can be written as $f(x) = x^\top Q x + r^\top x$, where

$$Q = \begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix}, \quad r = \begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix}.$$

(b) The function f is bounded below because the matrix Q is positive definite. This can be confirmed by checking that all eigenvalues of Q are positive (they are 1, 4, and 4).

(c) The unconstrained minimum occurs where $\nabla f(x) = 0$:

$$2Qx + r = 0 \implies x_u = -\frac{1}{2}Q^{-1}r = \begin{bmatrix} \frac{7}{8} \\ 1 \\ \frac{9}{8} \end{bmatrix}.$$

The minimum value is $f_u = -\frac{25}{8}$.

(d) With the constraint $c^\top x = b$, where $c = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$ and $b = 100$, we form the Lagrangian:

$$L(x, \nu) = x^\top Qx + r^\top x + \nu(c^\top x - b).$$

Setting the gradient to zero:

$$2Qx + r + \nu c = 0.$$

We solve this system along with the constraint $c^\top x = b$ to find x_c and ν. The system can be written as:

$$\begin{bmatrix} 2Q & c \\ c^\top & 0 \end{bmatrix} \begin{bmatrix} x \\ \nu \end{bmatrix} = \begin{bmatrix} -r \\ b \end{bmatrix}.$$

Solving the linear system, we obtain the unique optimal solution:

$$x_c = \begin{bmatrix} 3749/604 \\ 1997/302 \\ 4239/604 \end{bmatrix} \approx \begin{bmatrix} 6.2086 \\ 6.6160 \\ 7.0199 \end{bmatrix}, \quad \nu = -\frac{339}{151} \approx -2.244.$$

The constrained minimum value is $f_c = \frac{55573}{604} \approx 92.0083$.

(e) Geometrically, adding a constraint restricts the feasible set, so the constrained minimum f_c cannot be less than the unconstrained minimum f_u.

3. **KKT Equation in Terms of Jacobian**

The Jacobian matrix J is of size $p \times n$, where each row is $\nabla h_j(x)^\top$. The KKT condition $\nabla_x L = 0$ becomes:

$$\nabla f(x^*) + J^\top \vec{\nu} = 0.$$