Inequality-Constrained KKT System

Thomas R. Cameron

April 17, 2026

1 Introduction

Previously, we introduced inequality-constrained quadratic programs of the form

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & A\mathbf{x} \leq \mathbf{b} \end{aligned}$$

where $Q \in \mathbb{R}^{n\times n}$ is symmetric, $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, and $\mathbf{c} \in \mathbb{R}^n$. We defined the feasible region of this QP by

$$\mathcal{F} = \{\mathbf{x} \in \mathbb{R}^n \colon A\mathbf{x} \leq \mathbf{b}\}$$

and noted that this set is a polyhedron defined by the intersection of halfspaces.

Let $\mathbf{x} \in \mathcal{F}$. Then, the ith constrain is active at $\mathbf{x}$ if

$$\mathbf{a}_i^T\mathbf{x} = b_i$$

and inactive if $\mathbf{a}_i^T\mathbf{x} < b_i$. We denote by $\mathcal{A}(\mathbf{x})$ the set of indices corresponding to active constraints at $\mathbf{x}$. A vector $\mathbf{d} \in \mathbb{R}^n$ is called a feasible direction at $\mathbf{x}$ if there exists $\epsilon > 0$ such that

$$\mathbf{x} + t\mathbf{d} \in \mathcal{F},$$

for all $t \in [0, \epsilon]$. We saw that the feasible directions are described by the active constraints. In particular, $\mathbf{d}$ is a feasible direction at $\mathbf{x}$ if and only if

$$\mathbf{a}_i^T\mathbf{d} \leq 0,$$

for all $i \in \mathcal{A}(\mathbf{x})$. If $\mathbf{x}$ is an interior point, then there are no active constraints and all directions are feasible. The set of all feasible directions at $\mathbf{x} \in \mathcal{F}$ forms a convex cone:

$$\mathcal{C}_{\mathbf{x}} = \left\{\mathbf{d} \in \mathbb{R}^n \colon \mathbf{a}_i^T\mathbf{d} \leq 0, \quad \forall i \in \mathcal{A}(\mathbf{x})\right\}.$$

We saw that if $\mathbf{x}^* \in \mathcal{F}$ is optimal, then

$$\nabla f(\mathbf{x}^*)^T\mathbf{d} \geq 0,$$

for all $\mathbf{d} \in \mathcal{C}_{\mathbf{x}^*}$.

In this lecture, we use the feasible direction conditions and the necessary optimality conditions to form a KKT system for inequality-constrained QPs.

2 The KKT System

The KKT system encodes necessary conditions of feasibility and optimality. To construct this encoding for inequality-constrained QPs, we need heavier machinery known as Farkas lemma.

2.1 Farkas Lemma

Farkas lemma is a solvability theorem for a finite system of linear inequalities. It was originally proven by the Hungarian mathematician Gyula Farkas in 1902. Farkas lemma takes on many variations, but the one we will use is stated in the theorem below.

Theorem 2.1 (Farkas Lemma)**.** *Let $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. Then, exactly one of the two assertions is true*

(a) There exists a $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \leq 0$.

(b) There exists a $\mathbf{y} \in \mathbb{R}^m$ such that $A^T\mathbf{y} \leq 0$ and $\mathbf{b}^T\mathbf{y} < 0$.

Example

Let

$$A = \begin{bmatrix} 6 & 4 \\ 3 & 0 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The only solution to $A\mathbf{x} = \mathbf{b}$ is $x_1 = 1/3$ and $x_2 = -1/4$. Hence, the first assertion of Farkas Lemma does not hold. Therefore, there must exist a $\mathbf{y} \in \mathbb{R}^2$ such that $A^T\mathbf{y} \leq 0$ and $\mathbf{b}^T\mathbf{y} < 0$. Indeed, $y_1 = -1$ and $y_2 = 0$ satisfy.

Next, consider

$$A = \begin{bmatrix} 6 & 4 \\ 3 & 0 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} -6 \\ -3 \end{bmatrix}.$$

Now, the only solution to $A\mathbf{x} = \mathbf{b}$ is $x_1 = -1$ and $x_2 = 0$. So, the first assertion of Farkas Lemma holds; hence, the second assertion must be false. Indeed, $A^T\mathbf{y} \leq 0$ implies that $4y_1 \leq 0$ and $6y_1 + 3y_2 \leq 0$. However, $\mathbf{b}^T\mathbf{y} < 0$ implies that $6y_1 + 3y_2 > 0$, so we have a contradiction. □

2.2 First Order Optimality Condition

The first order optimality condition for inequality constrained QPs states that if $\mathbf{x}^*$ is optimal then

$$\nabla f(\mathbf{x}^*)^T\mathbf{d} \geq 0,$$

for all $\mathbf{d} \in \mathcal{C}_{\mathbf{x}^*}$ The following theorem states that this optimality condition is equivalent to $-\nabla f(\mathbf{x}^*)$ being in the cone generated by the active constraints.

Theorem 2.2. *Suppose that $\mathbf{x}^*$ is an optimal solution of the QP. Then, for each $i \in \mathcal{A}(\mathbf{x}^*)$ there exists a $y_i \geq 0$ such that*

$$-\nabla f(\mathbf{x}^*) = \sum_{i\in\mathcal{A}(\mathbf{x}^*)} y_i\mathbf{a}_i.$$

Proof. If $\mathbf{x}^*$ is an interior point, then $\mathcal{A}(\mathbf{x}^*)$ is empty. In this case, every vector $\mathbf{d} \in \mathbb{R}^n$ is a feasible direction. Hence,

$$\nabla f(\mathbf{x}^*)^T\mathbf{d} \geq 0,$$

for all $\mathbf{d} \in \mathbb{R}^n$, which implies that the gradient is zero and the result holds.

Suppose that $\mathbf{x}^*$ is a boundary point. Then, $\mathcal{A}(\mathbf{x}^*)$ is non-empty and the feasible directions are determined by the active constraints. Let $\hat{A}$ be the matrix whose ith column is $\mathbf{a}_i$ for each $i \in \mathcal{A}(\mathbf{x}^*)$. Then, every $\mathbf{d} \in \mathbb{R}^n$ that satisfies $\hat{A}^T\mathbf{d} \leq 0$ is a feasible direction. Since $\mathbf{x}^*$ is optimal, no vector $\mathbf{d} \in \mathbb{R}^n$ satisfies both $\hat{A}^T\mathbf{d} \leq 0$ and $\nabla f(\mathbf{x}^*)^T\mathbf{d} < 0$. Therefore, Farkas Lemma states that there must exist $\hat{\mathbf{y}} \leq 0$ such that

$$\nabla f(\mathbf{x}^*)^T = \hat{A}\hat{\mathbf{y}}$$

For each $i \in \mathcal{A}(\mathbf{x}^*)$, let $y_i = -\hat{y}_i$. Then, we have identified multipliers $y_i \geq 0$ such that

$$-\nabla f(\mathbf{x}^*) = \sum_{i \in \mathcal{A}(\mathbf{x}^*)} y_i \mathbf{a}_i.$$

□

The KKT system combines the first order optimality condition with feasibility constraints and slackness conditions. In particular, the KKT system for inequality-constrained QPs is defined below

$$Q\mathbf{x} + \mathbf{c} + A^T\mathbf{y} = 0, \tag{1}$$
$$A\mathbf{x} \leq \mathbf{b}, \tag{2}$$
$$\mathbf{y} \geq 0, \tag{3}$$
$$y_i\left(\mathbf{a}_i^T\mathbf{x} - b_i\right) = 0, \qquad \forall i \in \{1, \dots, m\}. \tag{4}$$

Note that (1) is the first order optimality condition by Theorem 2.2. Moreover, (2) is the primal feasibility conditions and (3) is the dual feasibility conditions. Finally, (4) describe the complementary slackness conditions. In particular, either the ith inequality is active or the ith multiplier is zero.

3 The Lagrangian and Interpretation of Multipliers

The Lagrangian associated with the inequality-constrained QP is defined as

$$L(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T\mathbf{x} + \mathbf{y}^T(A\mathbf{x} - \mathbf{b}),$$

where $\mathbf{y} \in \mathbb{R}^m$ satisfies $\mathbf{y} \geq \mathbf{0}$.

We can also write

$$L(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T\mathbf{x} + \sum_{i=1}^{m} y_i(\mathbf{a}_i^T\mathbf{x} - b_i).$$

The Lagrangian augments the objective function by penalizing constraint violations:

- If $\mathbf{a}_i^T\mathbf{x} > b_i$, then the term $y_i(\mathbf{a}_i^T\mathbf{x} - b_i)$ increases the value of L.
- If $\mathbf{a}_i^T\mathbf{x} \leq b_i$, the constraint is satisfied and contributes no penalty when $y_i = 0$.

The multipliers y_i control how strongly each constraint is enforced.

3.1 Recovering the KKT System

We compute the gradients of the Lagrangian:

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \mathbf{y}) = Q\mathbf{x} + \mathbf{c} + A^T \mathbf{y},$$
$$\nabla_{\mathbf{y}} L(\mathbf{x}, \mathbf{y}) = A\mathbf{x} - \mathbf{b}.$$

The KKT conditions can therefore be written compactly as:

$$\begin{aligned}
\nabla_{\mathbf{x}} L(\mathbf{x}, \mathbf{y}) &= 0 && \text{(first order optimality)} \\
\nabla_{\mathbf{y}} L(\mathbf{x}, \mathbf{y}) &\leq 0 && \text{(primal feasibility)} \\
\mathbf{y} &\geq \mathbf{0} && \text{(dual feasibility)} \\
\mathbf{y} \circ \nabla_{\mathbf{y}} L(\mathbf{x}, \mathbf{y}) &= 0 && \text{(complementary slackness)}
\end{aligned}$$

where the $\circ$ denotes the Hadamard (elementwise) product.

3.2 Weak and Strong Duality

We define the dual function

$$g(\mathbf{y}) = \inf_{\mathbf{x} \in \mathbb{R}^n} L(\mathbf{x}, \mathbf{y}),$$

where $\mathbf{y} \geq 0$ and the infimum inf denotes the greatest lower bound. The following theorem handles weak duality.

Theorem 3.1 (Weak Duality)**.** *Suppose that the QP is feasible. If $\overline{\mathbf{x}}$ is a feasible solution, then*

$$g(\mathbf{y}) \leq f(\overline{\mathbf{x}}),$$

for all $\mathbf{y} \geq 0$.

Proof. Suppose that $\overline{\mathbf{x}}$ is feasible. Then, $A\mathbf{x} \leq \mathbf{b}$, so $\mathbf{y}^T (A\mathbf{x} - \mathbf{b}) \leq 0$ for all $\mathbf{y} \geq 0$. Therefore,

$$L(\overline{\mathbf{x}}, \mathbf{y}) = f(\overline{\mathbf{x}}) + \mathbf{y}^T (A\mathbf{x} - \mathbf{b}) \leq f(\overline{\mathbf{x}}),$$

for all $\mathbf{y} \geq 0$. By definition of the infimum,

$$g(\mathbf{y}) = \inf_{\mathbf{x} \in \mathbb{R}^n} L(\mathbf{x}, \mathbf{y}) \leq L(\overline{\mathbf{x}}, \mathbf{y}) \leq f(\overline{\mathbf{x}}),$$

for all $\mathbf{y} \geq 0$. □

Thus every choice of multipliers gives a lower bound on the value of every primal feasible point. This is the first major reason the dual is important. It does not merely provide another optimization problem, it provides certificates in the form of lower bounds.

The dual problem is defined as follows

$$\sup_{\mathbf{y} \in \mathbb{R}^m} g(\mathbf{y}),$$

where the supremum sup is the least upper bound. Under certain conditions, this supremum (maximum) is attained at $\mathbf{y}^*$ and $g(\mathbf{y}^*) = f(\mathbf{x}^*)$, where $\mathbf{x}^*$ is an optimal solution to the QP. This is known as strong duality and is discussed in the following theorem.

Theorem 3.2 (Strong Duality)**.** *Suppose that the QP is feasible and Q is positive semidefinite. Then, for any solution $(\mathbf{x}^*, \mathbf{y}^*)$ to the KKT system, we have*

$$g(\mathbf{y}^*) = f(\mathbf{x}^*).$$

Moreover, $\mathbf{x}^*$ is an optimal solution to the QP.

Proof. Suppose that $(\mathbf{x}^*, \mathbf{y}^*)$ is a solution to the KKT system. Since $A\mathbf{x}^* \leq \mathbf{b}$, we have

$$L(\mathbf{x}^*, \mathbf{y}) \leq f(\mathbf{x}^*),$$

for all $\mathbf{y} \geq 0$. In particular, weak duality states that $g(\mathbf{y}) \leq f(\mathbf{x}^*)$, for all $\mathbf{y} \geq 0$.

Now, the gradient of $L(\mathbf{x}, \mathbf{y}^*)$ with respect to $\mathbf{x}$ is given by

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \mathbf{y}^*) = Q\mathbf{x} + \mathbf{c} + A^T\mathbf{y}^*.$$

Hence, $\mathbf{x}^*$ is a critical point of $L(\mathbf{x}, \mathbf{y}^*)$. Moreover, the Hessian of $L(\mathbf{x}, \mathbf{y}^*)$, with respect to $\mathbf{x}$, is given by

$$\nabla^2_{\mathbf{xx}} L(\mathbf{x}, \mathbf{y}^*) = Q,$$

which is positive semidefinite. Therefore, $L(\mathbf{x}, \mathbf{y}^*)$ is convex and the infimum (minimum) is attained at $\mathbf{x}^*$. Furthermore,

$$g(\mathbf{y}^*) = \inf_{\mathbf{x} \in \mathbb{R}^n} L(\mathbf{x}, \mathbf{y}^*) = L(\mathbf{x}^*, \mathbf{y}^*) \leq f(\mathbf{x}^*).$$

Now, we apply the complementary slackness conditions, which state that

$$y_i^* \left(\mathbf{a}_i^T\mathbf{x}^* - b_i\right) = 0,$$

for all $i = 1, \ldots, m$. Hence,

$$g(\mathbf{y}^*) = \inf_{\mathbf{x} \in \mathbb{R}^n} L(\mathbf{x}, \mathbf{y}^*) = L(\mathbf{x}^*, \mathbf{y}^*) = f(\mathbf{x}^*).$$

□

For equality-constrained QPs, we were able to prove that the KKT system is feasible if and only if $g(\mathbf{y})$ is finite. For inequality-constrained QPs, we no longer have this connection. Instead, the convexity of Q combined with Slater's condition guarantees that the KKT system is feasible.

4 The Active-Set Method

The KKT conditions tell us that at an optimal solution, some inequality constraints are active and the rest are inactive. If we knew in advance which constraints were active, then we could solve the problem as an equality-constrained QP. This observation motivates the active-set method.

The idea is to maintain a set W of constraints, called the *working set*, that are treated as equalities. At each step, we solve the equality-constrained QP determined by W. Then we check two things:

- whether the resulting point is feasible for all inequality constraints, and
- whether the corresponding multipliers are nonnegative.

If a constraint is violated, then it should be added to the working set. If a multiplier is negative, then the corresponding constraint should be removed from the working set. The process continues until the KKT conditions are satisfied.

4.0.1 Example

Consider the QP

$$\begin{aligned} \text{minimize} \quad & x_1^2 + x_1 x_2 + 2x_2^2 - 4x_1 - 6x_2 \\ \text{subject to} \quad & x_1 + x_2 \le 2, \\ & x_2 \le \frac{3}{2}, \\ & x_1, x_2 \ge 0. \end{aligned}$$

Suppose that we start with a working set $W = \emptyset$. Then, we solve the unconstrained problem

$$\text{minimize} \quad f(\mathbf{x}) = x_1^2 + x_1 x_2 + 2x_2^2 - 4x_1 - 6x_2$$

which has a minimum when $\nabla f(\mathbf{x}) = 0$, that is, when

$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}.$$

The unique solution to this system is

$$\mathbf{x}^* = \frac{1}{7} \begin{bmatrix} 10 \\ 8 \end{bmatrix}.$$

Note that this solution does not satisfy the first constraint since

$$\frac{10}{7} + \frac{8}{7} = \frac{18}{7} > 2.$$

Next, we update the working set by adding the violated constraint, so $W = \{1\}$. Then, we solved the constrained problem

$$\begin{aligned} \text{minimize} \quad & x_1^2 + x_1 x_2 + 2x_2^2 - 4x_1 - 6x_2 \\ \text{subject to} \quad & x_1 + x_2 = 2 \end{aligned}$$

The corresponding KKT system is

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ y_1 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ 2 \end{bmatrix},$$

which has a solution of

$$x_1 = 1, \quad x_2 = 1, \quad y_1 = 1.$$

Note that the Q is positive definite, so this KKT solution is an optimal solution to the equality-constrained subproblem. Moreover, this solution satisfies all constraines of the inequality-constrained QP and is therefore an optimal solution to the original problem.