Lecture 13: KKT Examples Second-Order Optimality

CS 169 / CS 268
Introduction to Optimization

Ioannis Panageas

November 12, 2026

KKT recap

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) \quad \text{s.t.} \quad g_i(\boldsymbol{x}) \leq 0, \quad h_j(\boldsymbol{x}) = 0.$$

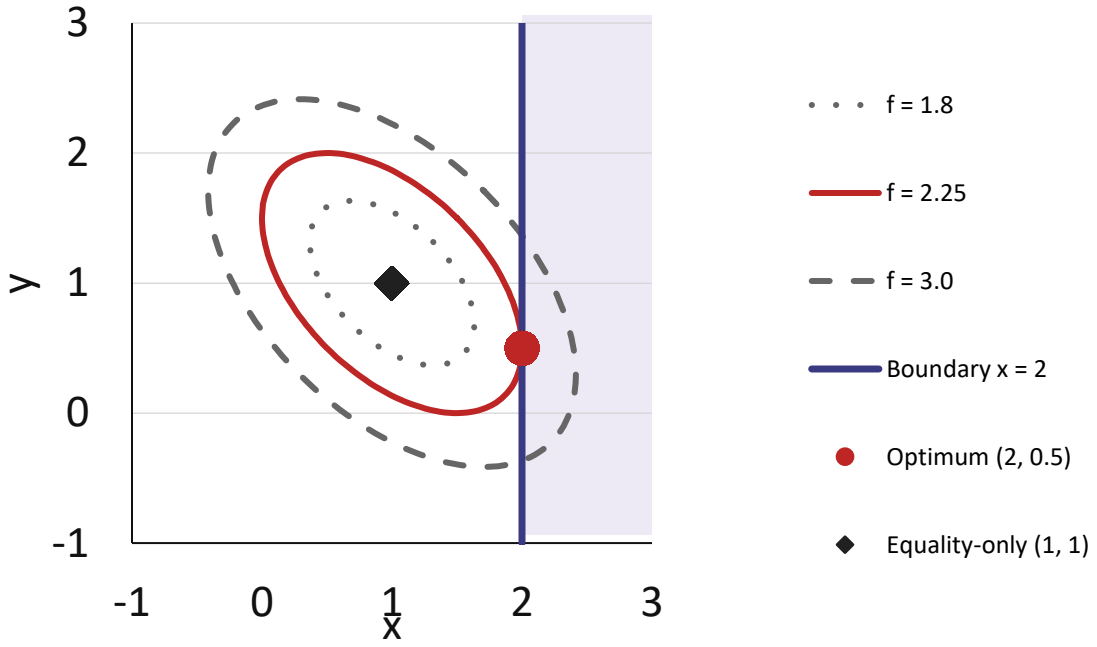
$$\mathcal{L}(\boldsymbol{x}, \lambda, \nu) = f(\boldsymbol{x}) + \sum_i \lambda_i g_i(\boldsymbol{x}) + \sum_j \nu_j h_j(\boldsymbol{x}).$$

$$\begin{aligned}
\nabla_x \mathcal{L}(\boldsymbol{x}^*, \lambda^*, \nu^*) = 0 \quad & \text{stationarity,} \\
g_i(\boldsymbol{x}^*) \leq 0, \quad h_j(\boldsymbol{x}^*) = 0 \quad & \text{primal feasibility,} \\
\lambda_i^* \geq 0 \quad & \text{dual feasibility,} \\
\lambda_i^* g_i(\boldsymbol{x}^*) = 0 \quad & \text{complementary slackness.}
\end{aligned}$$

Equality multipliers ν_j can have either sign.

Example 1: a plane and a bound

$$\min_{x,y,z} \frac{1}{2}(x^2+y^2+z^2) \quad \text{s.t.} \quad \begin{cases} x+y+z=3, \\ x \geq 2. \end{cases}$$



Eliminate $z = 3 - x - y$. The equality-only minimum $(1,1,1)$ violates $x \geq 2$.

Example 1: the KKT equations

$$\mathcal{L} = \frac{1}{2}(x^2 + y^2 + z^2) + \lambda(2 - x) + \nu(x + y + z - 3).$$

$$\begin{aligned} x - \lambda + \nu &= 0, \\ y + \nu &= 0, \\ z + \nu &= 0. \end{aligned} \qquad \begin{aligned} x + y + z &= 3, \\ x &\geq 2, \\ \lambda &\geq 0, \\ \lambda(2 - x) &= 0. \end{aligned}$$

If $\lambda = 0$, stationarity gives $x = y = z = 1$: infeasible. Hence $\lambda > 0$ and $x = 2$.

Example 1: solution and certificate

$$x = 2, \qquad y = z, \qquad y + z = 1.$$

$$\boxed{(x^*, y^*, z^*) = (2, \tfrac{1}{2}, \tfrac{1}{2}), \quad \lambda^* = \tfrac{3}{2}, \quad \nu^* = -\tfrac{1}{2}.}$$

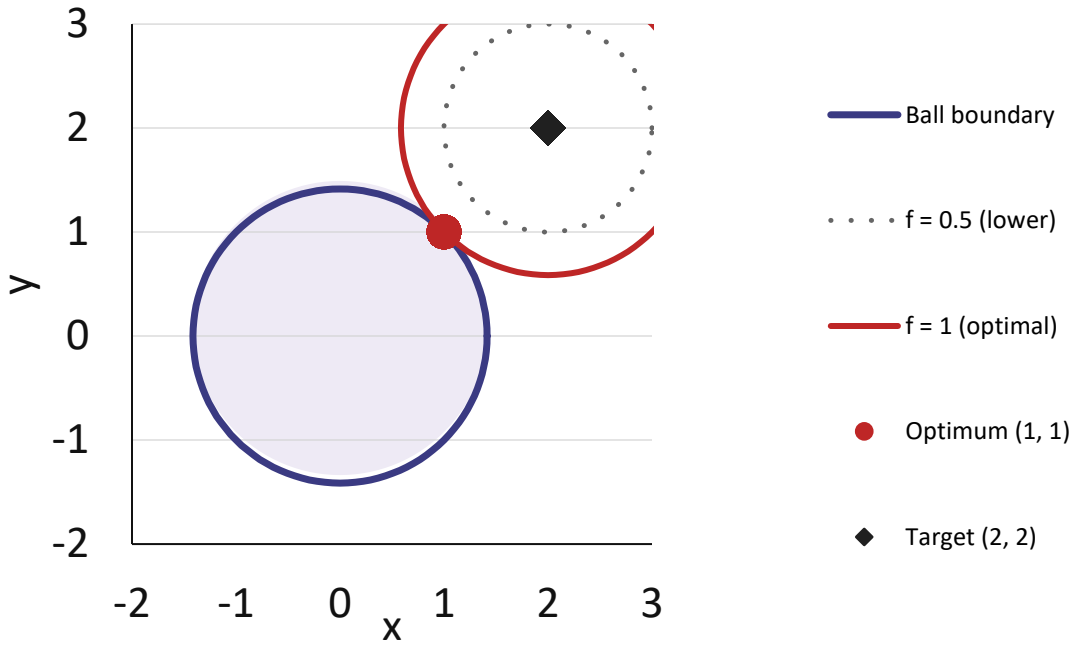
A direct check: use $y^2 + z^2 \geq (y+z)^2/2$ and $y + z = 3 - x$.

$$f(x,y,z) \geq \frac{x^2}{2} + \frac{(3-x)^2}{4} = \frac{9}{4} + \frac{3}{4}(x-2)^2 + \frac{3}{2}(x-2) \geq \frac{9}{4}.$$

Equality holds only at $(2, \frac{1}{2}, \frac{1}{2})$: the unique global minimum.

Example 2: a ball and a bound

$$\min_{x,y,z} \ \frac{1}{2}\left((x-2)^2+(y-2)^2+z^2\right) \quad \text{s.t.} \quad \begin{cases} x^2+y^2+z^2 \le 2, \\ z \le 1. \end{cases}$$



The figure shows the slice $z = 0$. We will derive $z^* = 0$ from KKT.

Example 2: the KKT equations

$$\mathcal{L} = \frac{1}{2}\left((x-2)^2 + (y-2)^2 + z^2\right) + \lambda(x^2 + y^2 + z^2 - 2) + \mu(z-1).$$

$$(1+2\lambda)x = 2,$$
$$(1+2\lambda)y = 2,$$
$$(1+2\lambda)z + \mu = 0.$$

$$\lambda, \mu \geq 0, \qquad \lambda(x^2 + y^2 + z^2 - 2) = 0, \qquad \mu(z-1) = 0.$$

Also impose $x^2 + y^2 + z^2 \leq 2$ and $z \leq 1$.

Example 2: active and inactive

$$z = -\frac{\mu}{1+2\lambda} \leq 0 < 1 \quad \Longrightarrow \quad \mu = 0, \quad z = 0.$$

If $\lambda = 0$, then $x = y = 2$, which violates the ball constraint. Thus $\lambda > 0$ and the ball is active.

$$x = y = \frac{2}{1+2\lambda}, \qquad 2\left(\frac{2}{1+2\lambda}\right)^2 = 2.$$

$$\boxed{(x^*, y^*, z^*) = (1,1,0), \quad \lambda^* = \tfrac{1}{2}, \quad \mu^* = 0, \quad f^* = 1.}$$

Convex KKT sufficiency

Assume f, g_i are differentiable and convex, and h_j are affine. Suppose (x^*, λ^*, ν^*) satisfies KKT.

Fix the multipliers. Since $\lambda_i^* \geq 0$, the function $\mathcal{L}(x, \lambda^*, \nu^*)$ is convex in x.

$$\nabla_x \mathcal{L}(x^*, \lambda^*, \nu^*) = 0 \quad \Longrightarrow \quad x^* \in \arg\min_x \mathcal{L}(x, \lambda^*, \nu^*).$$

We can now compare x^* with every feasible point.

Step 1: equality at the KKT point

Keep (λ^*, ν^*) fixed throughout the proof. First evaluate the Lagrangian at the KKT point $\boldsymbol{x}^*$.

$$\mathcal{L}(\boldsymbol{x}^*, \lambda^*, \nu^*) = f(\boldsymbol{x}^*) + \sum_i \lambda_i^* g_i(\boldsymbol{x}^*) + \sum_j \nu_j^* h_j(\boldsymbol{x}^*).$$

Feasibility: $h_j(\boldsymbol{x}^*) = 0$ for every j.

Complementary slackness: $\lambda_i^* g_i(\boldsymbol{x}^*) = 0$ for every i.

$$\boxed{\mathcal{L}(\boldsymbol{x}^*, \lambda^*, \nu^*) = f(\boldsymbol{x}^*) + 0 + 0 = f(\boldsymbol{x}^*).}$$

Step 2: the first inequality

The fixed-multiplier Lagrangian is convex in $\boldsymbol{x}$: f, g_i are convex, $\lambda_i^* \geq 0$, and h_j are affine.

$$\mathcal{L}(\boldsymbol{u}, \lambda^*, \nu^*) \geq \mathcal{L}(\boldsymbol{x}^*, \lambda^*, \nu^*) + \left[\nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}^*, \lambda^*, \nu^*)\right]^T (\boldsymbol{u} - \boldsymbol{x}^*).$$

Stationarity gives $\nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}^*, \lambda^*, \nu^*) = 0$, so the inner-product term vanishes.

$$\boxed{\mathcal{L}(\boldsymbol{x}^*, \lambda^*, \nu^*) \leq \mathcal{L}(\boldsymbol{u}, \lambda^*, \nu^*).}$$

Step 3: comparison with feasible points

Now choose any feasible $\boldsymbol{u}$. Feasibility and $\lambda_i^* \geq 0$ give

$$\lambda_i^* g_i(\boldsymbol{u}) \leq 0, \qquad \nu_j^* h_j(\boldsymbol{u}) = 0.$$

$$\mathcal{L}(\boldsymbol{u}, \lambda^*, \nu^*) = f(\boldsymbol{u}) + \sum_i \lambda_i^* g_i(\boldsymbol{u}) \leq f(\boldsymbol{u}).$$

$$\boxed{f(\boldsymbol{x}^*) = \mathcal{L}(\boldsymbol{x}^*, \lambda^*, \nu^*) \leq \mathcal{L}(\boldsymbol{u}, \lambda^*, \nu^*) \leq f(\boldsymbol{u}).}$$

Since $\boldsymbol{u}$ was arbitrary, $\boldsymbol{x}^*$ is a global minimizer.

Strict convexity and uniqueness

If f is strictly convex and the feasible set is convex, there is at most one minimizer.

If distinct feasible points $\boldsymbol{u}, \boldsymbol{v}$ both attained the minimum f^*, their midpoint would be feasible and satisfy

$$f\left(\frac{\boldsymbol{u}+\boldsymbol{v}}{2}\right) < \frac{f(\boldsymbol{u})+f(\boldsymbol{v})}{2} = f^*.$$

Both recap objectives have Hessian I_3. Their KKT solutions are therefore the unique global minima.

Taylor expansion at a stationary point

Now allow nonconvex problems. Suppose f is twice continuously differentiable near $\boldsymbol{x}_0$ and $\nabla f(\boldsymbol{x}_0) = 0$.

$$f(\boldsymbol{x}_0 + \boldsymbol{h}) = f(\boldsymbol{x}_0) + \frac{1}{2}\boldsymbol{h}^T \nabla^2 f(\boldsymbol{x}_0)\boldsymbol{h} + o(\|\boldsymbol{h}\|^2).$$

$$\boldsymbol{h} = t\boldsymbol{d} \quad \Longrightarrow \quad f(\boldsymbol{x}_0 + t\boldsymbol{d}) - f(\boldsymbol{x}_0) = \frac{t^2}{2}\boldsymbol{d}^T \nabla^2 f(\boldsymbol{x}_0)\boldsymbol{d} + o(t^2).$$

The sign of $\boldsymbol{d}^T \nabla^2 f(\boldsymbol{x}_0)\boldsymbol{d}$ describes curvature in direction $\boldsymbol{d}$.

Unconstrained second-order tests

At a stationary point $\boldsymbol{x}_0$, write $H = \nabla^2 f(\boldsymbol{x}_0)$ and assume f is twice continuously differentiable.

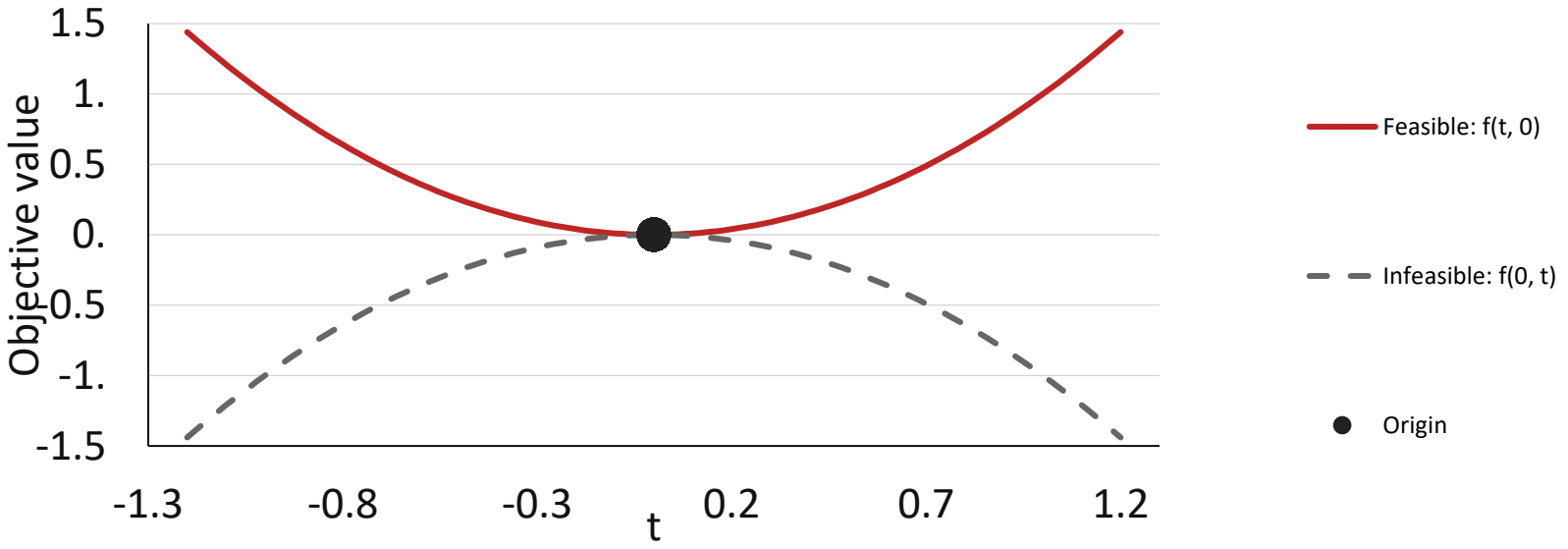
$$\boldsymbol{x}_0 \text{ local minimum} \quad \Longrightarrow \quad \boldsymbol{d}^T H \boldsymbol{d} \geq 0 \quad \text{for every } \boldsymbol{d}.$$

$$\boldsymbol{d}^T H \boldsymbol{d} > 0 \quad \text{for every } \boldsymbol{d} \neq 0 \quad \Longrightarrow \quad \boldsymbol{x}_0 \text{ strict local minimum.}$$

If some direction has negative curvature, the point is not a minimum. Zero curvature can be inconclusive: t^4 and $-t^4$ at 0.

Constraints remove some directions

$$\min_{x,y} \; x^2 - y^2 \quad \text{s.t.} \quad y = 0, \qquad \nabla^2 f = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$



$$f(x,0) = x^2 \geq 0, \qquad d_y = 0 \quad \Longrightarrow \quad \boldsymbol{d}^T \nabla^2 f \, \boldsymbol{d} = 2d_x^2.$$

Tangent directions for equalities

Consider $h_1(\boldsymbol{x}) = \cdots = h_m(\boldsymbol{x}) = 0$. A feasible curve $\boldsymbol{\gamma}(t)$ satisfies $h_j(\boldsymbol{\gamma}(t)) = 0$.

$$\boldsymbol{\gamma}(0) = \boldsymbol{x}^*, \quad \boldsymbol{\gamma}'(0) = \boldsymbol{d} \quad \Longrightarrow \quad \nabla h_j(\boldsymbol{x}^*)^T \boldsymbol{d} = 0.$$

$$\boxed{T = \{\boldsymbol{d} : \nabla h_j(\boldsymbol{x}^*)^T \boldsymbol{d} = 0 \text{ for every } j\}.}$$

Assume the constraints are twice continuously differentiable and their gradients at $\boldsymbol{x}^*$ are linearly independent. Then T describes the initial directions of motion along the constraints.

The Hessian of the Lagrangian

$$\mathcal{L}(x, \nu) = f(x) + \sum_{j=1}^{m} \nu_j h_j(x).$$

At a KKT point, keep the multipliers ν^* fixed and differentiate twice with respect to x:

$$\boxed{H = \nabla^2_{xx} \mathcal{L}(x^*, \nu^*) = \nabla^2 f(x^*) + \sum_j \nu_j^* \nabla^2 h_j(x^*).}$$

The constraint Hessians account for the bending of feasible curves. For affine equalities, they vanish and $H = \nabla^2 f$.

Curvature on the feasible set

Keep the multipliers ν^* fixed. At every feasible point, the equality terms vanish:

$$h_j(\boldsymbol{x}) = 0 \text{ for every } j \quad \Longrightarrow \quad \boxed{\mathcal{L}(\boldsymbol{x}, \nu^*) = f(\boldsymbol{x}).}$$

At a KKT point, $\nabla_x \mathcal{L} = 0$. The Hessian H describes the local curvature of this stationary function.

For the constrained problem, we examine $\boldsymbol{d}^T H \boldsymbol{d}$ only for tangent directions $\boldsymbol{d} \in T$.

Second-order conditions: equalities

Assume f, h_j are twice continuously differentiable, $\boldsymbol{x}^*$ is feasible, and the constraint gradients are linearly independent there.

Use $T = \{\boldsymbol{d} : \nabla h_j(\boldsymbol{x}^*)^T \boldsymbol{d} = 0 \; \forall j\}$ and $H = \nabla^2_{\boldsymbol{xx}} \mathcal{L}(\boldsymbol{x}^*, \nu^*)$.

Necessary: a local minimum has KKT multipliers satisfying

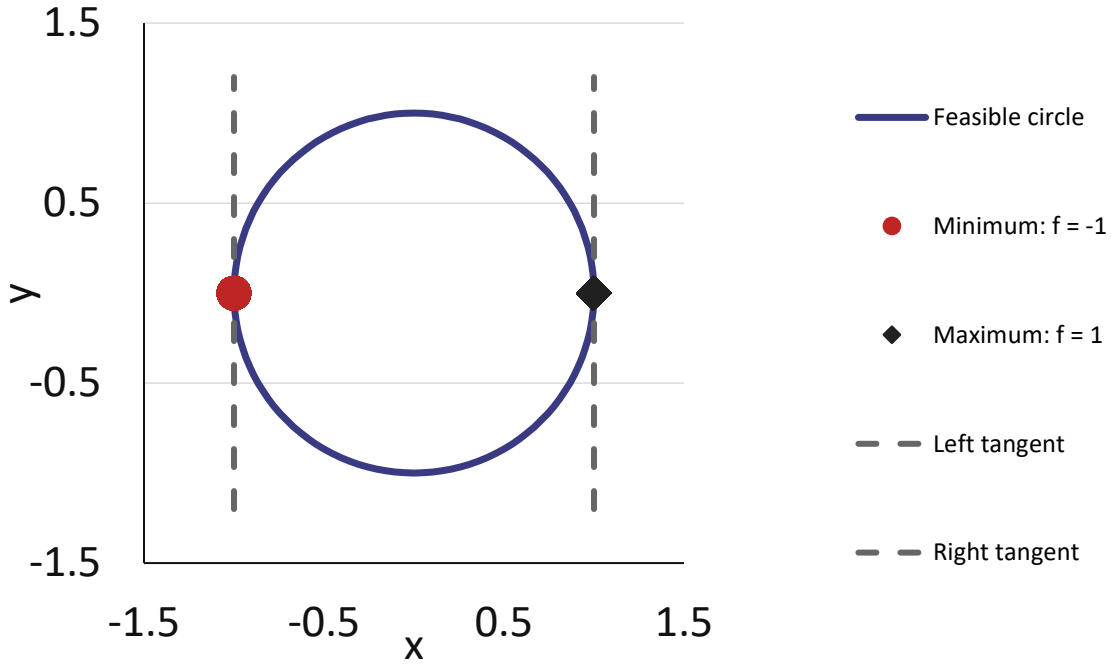
$$\boldsymbol{d}^T H \boldsymbol{d} \geq 0 \qquad \text{for every } \boldsymbol{d} \in T.$$

Sufficient: a KKT point is a strict local minimum if

$$\boldsymbol{d}^T H \boldsymbol{d} > 0 \qquad \text{for every nonzero } \boldsymbol{d} \in T.$$

Example: minimizing x on a circle

$$\min_{x,y} \; x \qquad \text{s.t.} \quad x^2 + y^2 = 1.$$



The feasible set is nonconvex. Both the minimum and maximum satisfy the first-order equations.

The two KKT points on the circle

$$\mathcal{L}(x, y, \nu) = x + \nu(x^2 + y^2 - 1).$$

$$1 + 2\nu x = 0, \qquad 2\nu y = 0, \qquad x^2 + y^2 = 1.$$

The first equation gives $\nu \neq 0$. Hence $y = 0$, and feasibility gives $x = \pm 1$.

$$(x, y) = (-1, 0): \quad \nu = \tfrac{1}{2},$$
$$(x, y) = (1, 0): \quad \nu = -\tfrac{1}{2}.$$

What changes with inequalities?

At an active constraint $g(\boldsymbol{x}^*) = 0$, a one-sided feasible curve with initial velocity $\boldsymbol{d}$ must satisfy

$$\nabla g(\boldsymbol{x}^*)^T \boldsymbol{d} \leq 0.$$

$x \geq 0$ at $x^* = 0$: $\quad g(x) = -x, \quad -d \leq 0 \iff d \geq 0.$