Constrained nonlinear optimization

Final class · November 16, 2026

2105623 · Optimization of Chemical Processes

From feasible steps to a verified solution

Plan for the final class

0–15 min · retrieve unconstrained NLP and line search

15–90 min · equality constraints and KKT

90–100 min · break; 100–155 min · methods and Python

155–180 min · synthesis, assignment and practice

Gradient descent needs a step rule

$f(u,v)=(u-1)^2+4(v+2)^2$; $\nabla f=(2(u-1),8(v+2))$

$r_{k+1}=r_k+\alpha_k p_k$; gradient descent uses $p_k=-\nabla f(r_k)$

At $r_0=(3,0)$: $g_0=(4,16)$, $g_0^T p_0=-272$

Armijo checks sufficient decrease before moving

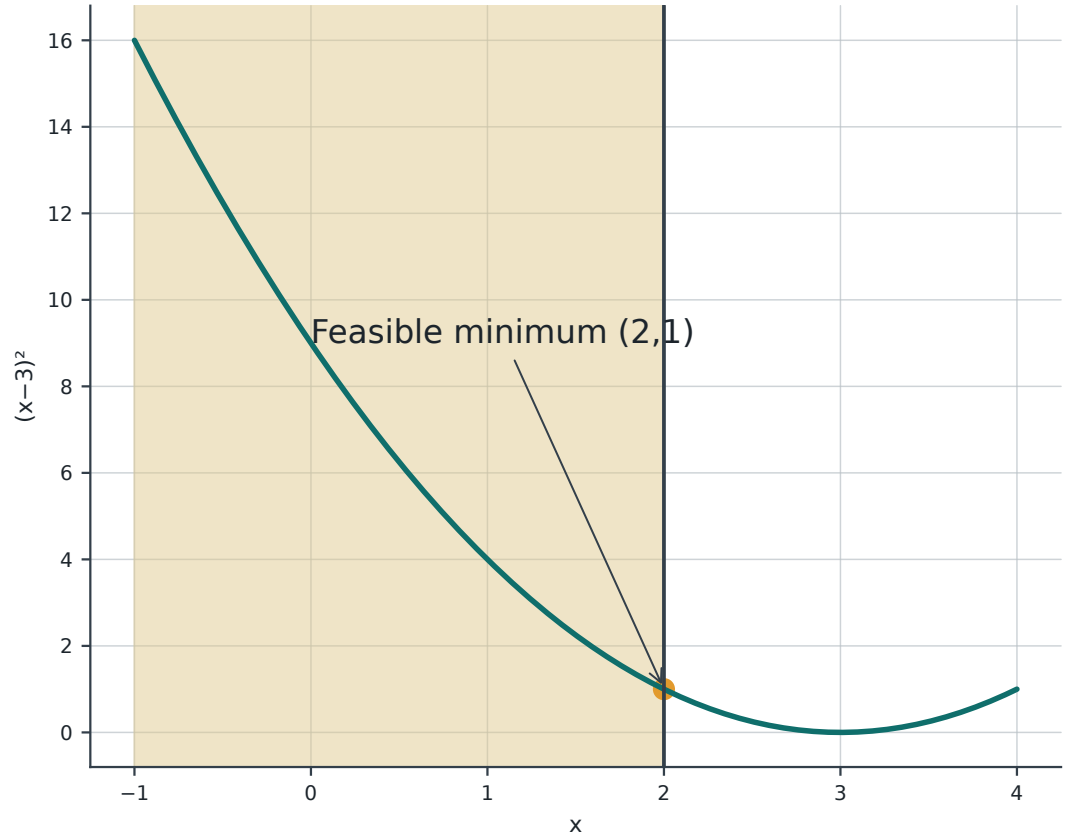
A good direction can leave the feasible region

min $f(x)=(x-3)^2$ subject to $x \le 2$

At $x=2$: $f'(2)=2(2-3)=-2$

Descent needs $d>0$, while feasibility needs $d \le 0$

The constrained minimum is $x^*=2$, with $f^*=1$



Write a consistent constrained model

min f(r), subject to $g_i(r) \leq 0$ and $h_j(r)=0$

$L(r,\lambda,\nu)=f(r)+\Sigma\lambda_i g_i(r)+\Sigma\nu_j h_j(r)$

$\lambda_i \geq 0$; equality multipliers ν_j are unrestricted

Convert maximization to minimization before using this convention

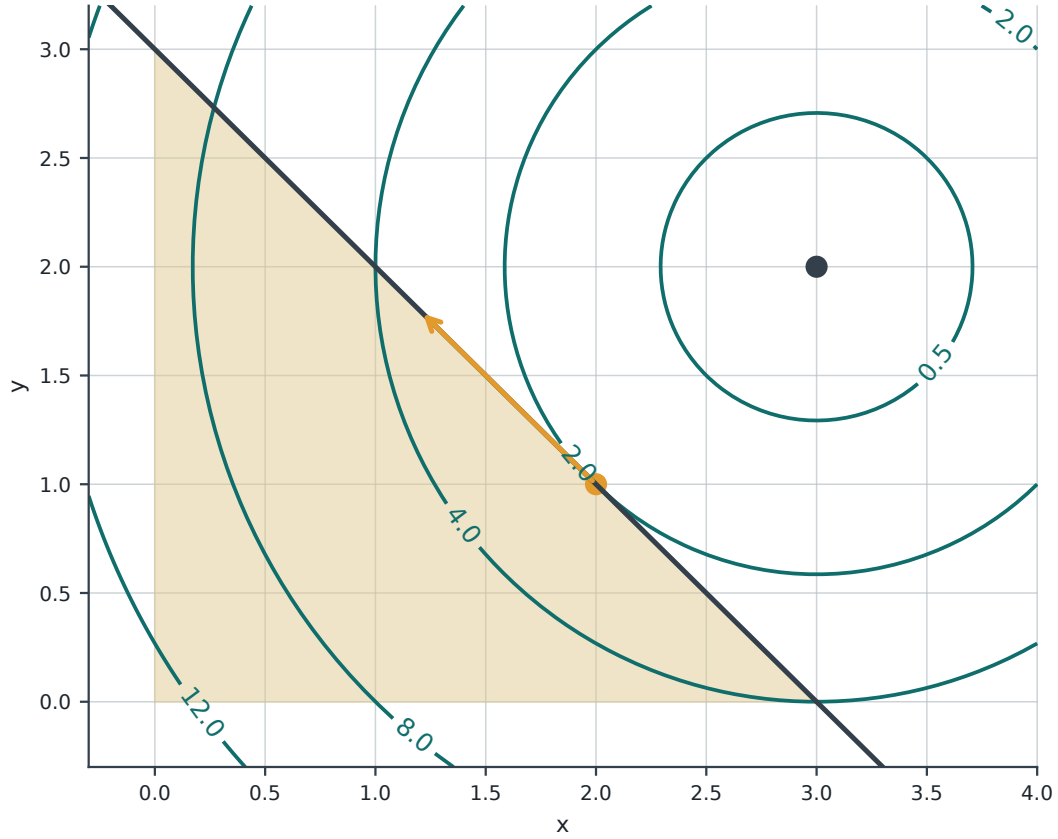
Feasible directions near an active boundary

If g(r)=0, a linearized feasible direction satisfies $\nabla g(r)^T d \leq 0$

For h(r)=0, the tangent condition is $\nabla h(r)^T d = 0$

For nonlinear constraints these are first-order conditions

A finite trial step still needs a feasibility check



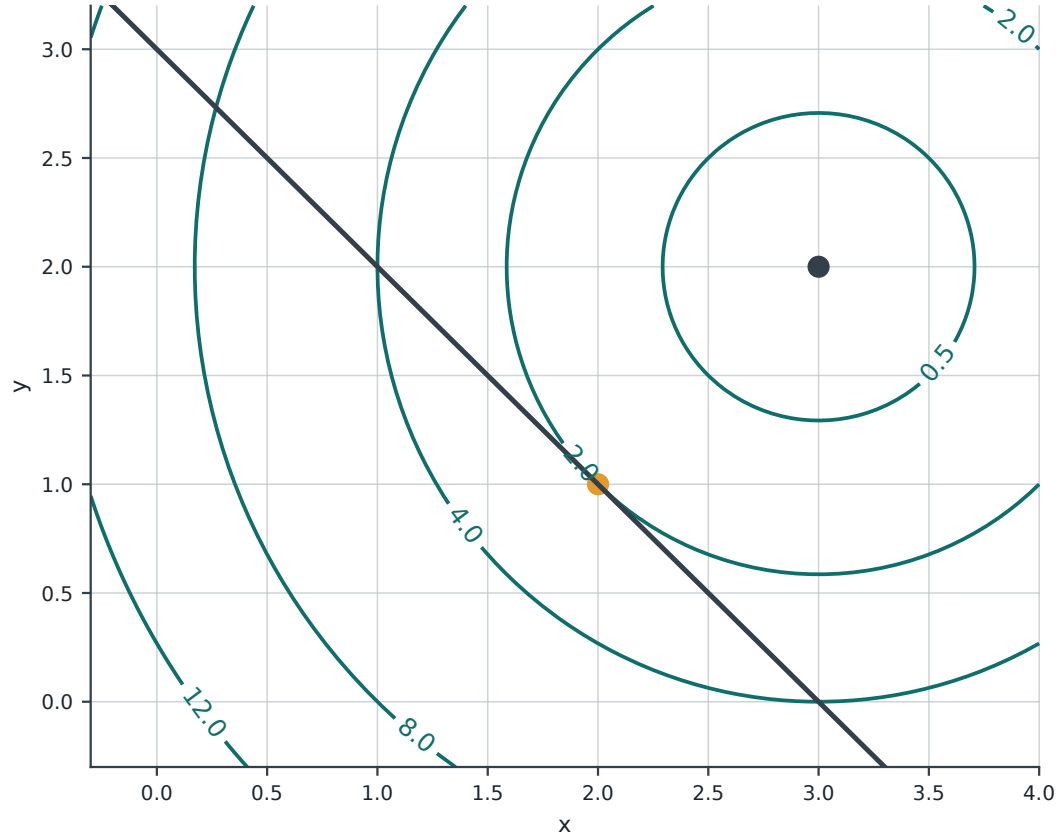
Equality constraints remove normal motion

$\min (x-3)^2+(y-2)^2$ subject to $x+y=3$

A tangent direction is $d=(1,-1)$

At a stationary feasible point, $\nabla f^T d=0$

Thus ∇f is normal to the constraint line



Lagrange equations for the equality example

$L=(x-3)^2+(y-2)^2+\nu(x+y-3)$

$2(x-3)+\nu=0;\ 2(y-2)+\nu=0$

$x+y=3$ gives $\nu=2$ and $(x,y)=(2,1)$

$f^*=2$; the objective is strictly convex

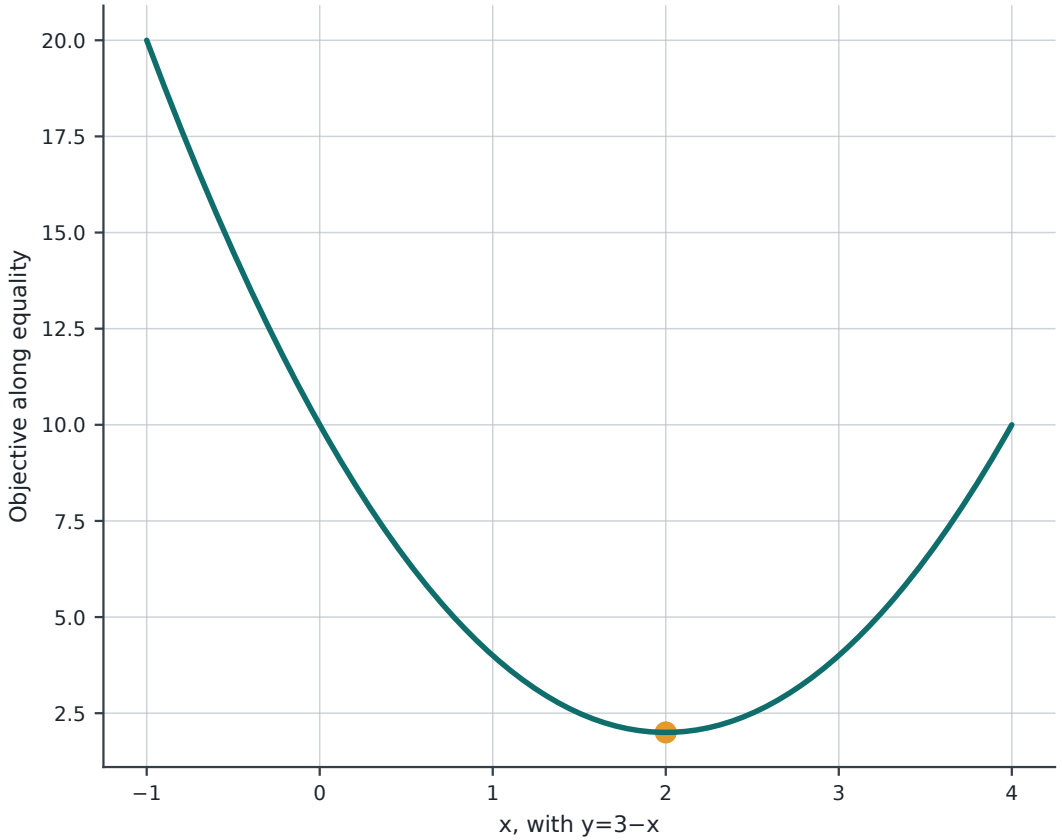
Check the answer by substitution

Let $y=3-x$

$f(x,3-x)=(x-3)^2+(1-x)^2$

Derivative $4x-8=0$ gives $x=2$; $y=1$

Second derivative is $4>0$, so this is the unique minimum



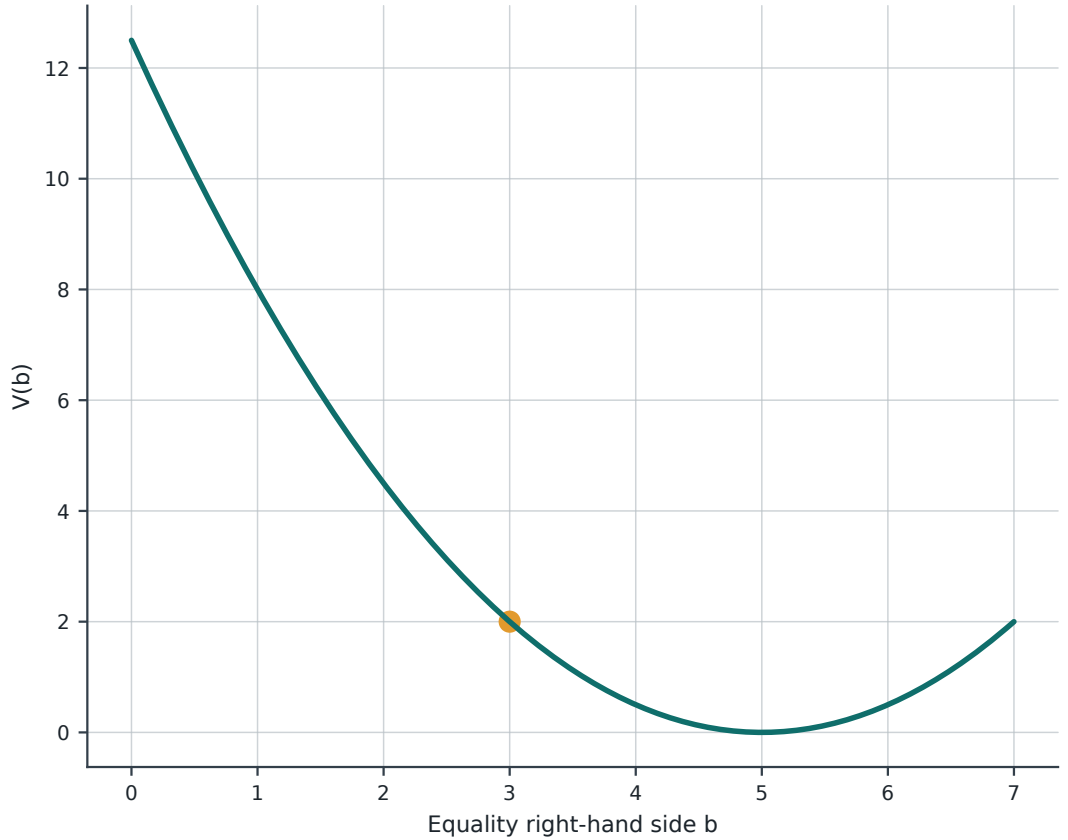
An equality multiplier measures a local change

Replace $x+y=3$ by $x+y=b$

For this quadratic: $x=(b+1)/2$, $y=(b-1)/2$

$V(b)=(b-5)^2/2$ and $\nu=5-b$

$dV/db=-\nu$; at $b=3$ the derivative is -2



KKT has four parts

Stationarity: $\nabla f + \Sigma \lambda_i \nabla g_i + \Sigma \nu_j \nabla h_j = 0$

Primal feasibility: $g_i \leq 0$ and $h_j = 0$

Dual feasibility: $\lambda_i \geq 0$

Complementary slackness: $\lambda_i g_i = 0$ for every i

Primal feasibility checks the proposed point

Substitute the candidate into every original constraint

Bounds are constraints too

Small objective value cannot repair an infeasible point

Report the largest violation and its units

Dual feasibility checks multiplier signs

For minimization with $g \le 0$ and $L = f + \lambda g$, require $\lambda \ge 0$

In min f subject to $a^T r \le b$, $dV/db = -\lambda$ locally

A tighter upper bound cannot improve the global minimum value

A negative λ rejects a candidate under this convention

Complementary slackness pairs price with slack

If $g_i<0$, then $\lambda_i=0$

If $\lambda_i>0$, then $g_i=0$

An active constraint can have $\lambda_i=0$

Use min x^2 subject to $x\leq 0$ at $x^*=0$ as a counterexample

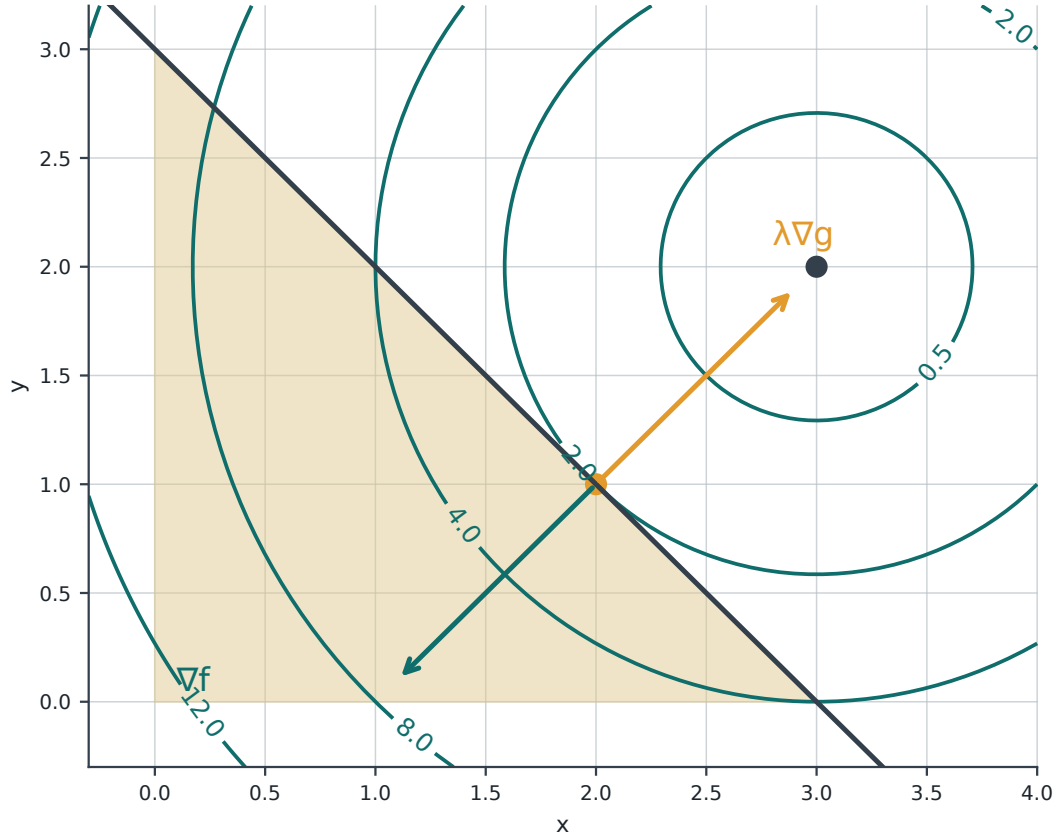
Stationarity balances objective and constraints

At a constrained optimum, ∇f may be nonzero

Constraint normals balance the objective gradient

At (2,1), ∇f=(−2,−2) and λ∇g=(2,2)

Their sum is zero when λ=2



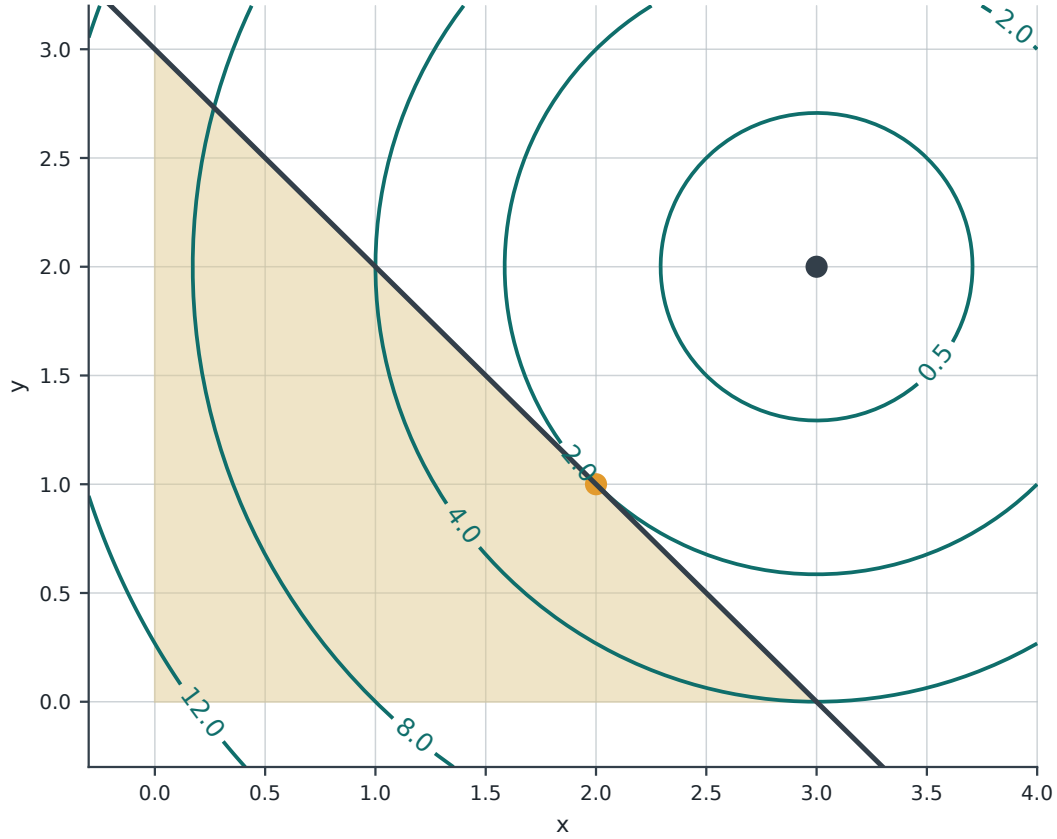
Inequality example with nonnegative production

$\min f=(x-3)^2+(y-2)^2$

$g_1=x+y-3\le 0;\ g_2=-x\le 0;\ g_3=-y\le 0$

Unconstrained minimum (3,2) violates capacity

We must find the active constraints



Test the capacity-active case

Assume $x+y=3$ and $x,y>0$, so $\lambda_2=\lambda_3=0$

$2(x-3)+\lambda_1=0$; $2(y-2)+\lambda_1=0$

Solve: $\lambda_1=2$, $x=2$, $y=1$

The assumed positive coordinates are consistent

Verify all KKT conditions

Primal: $(g_1, g_2, g_3) = (0, -2, -1)$

Dual: $(\lambda_1, \lambda_2, \lambda_3) = (2, 0, 0)$

Complementarity: every $\lambda_i g_i = 0$

Stationarity: $(-2, -2) + 2(1, 1) = (0, 0)$

Why this candidate is a global minimum

Hessian of f is 2I, which is positive definite

All constraints are affine; the feasible set is convex

A feasible KKT point is globally optimal for this convex problem

Strict convexity gives the unique optimum (2,1), f*=2

KKT necessity needs regularity

For differentiable constrained NLP, suitable constraint qualifications support KKT necessity

LICQ requires independent gradients of equality and active inequality constraints

For convex problems, Slater can support multiplier existence under its standard assumptions

For sufficiency in a convex problem, a verified KKT point is enough

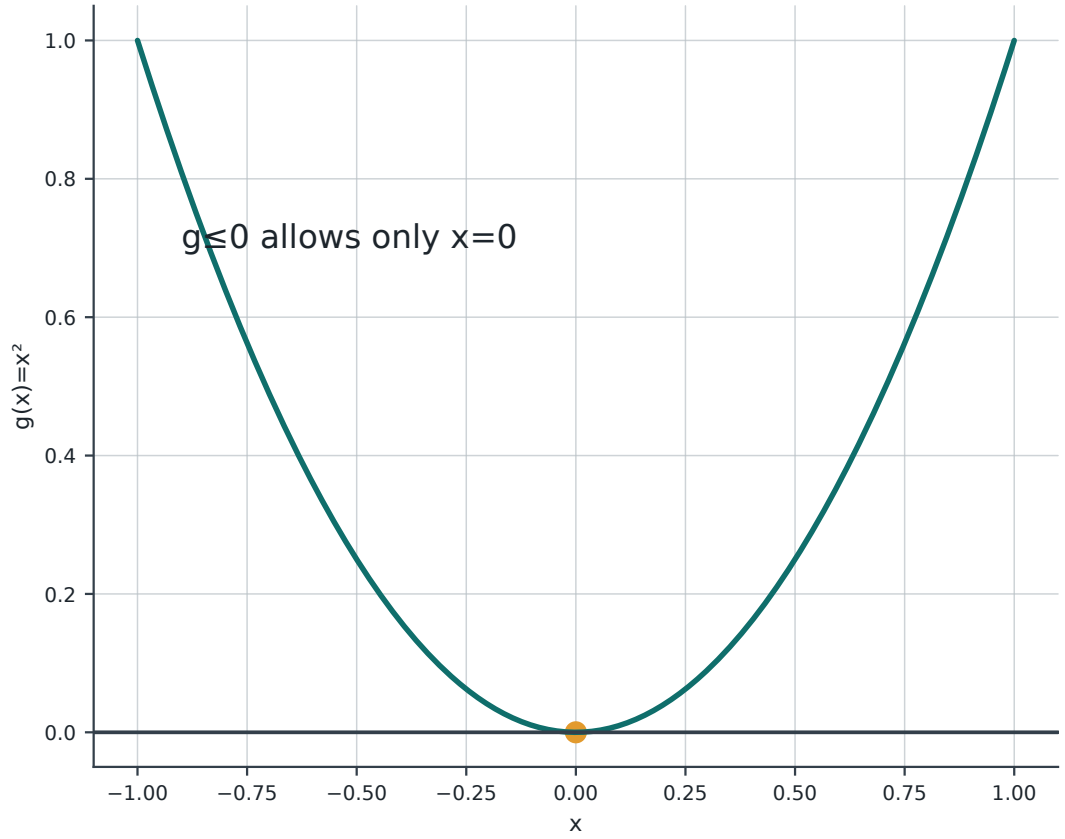
A minimum can fail KKT without regularity

min x subject to $x^2 \leq 0$

The only feasible point is $x^*=0$, so it is the minimum

Stationarity would require $1+\lambda(2x)=0$

At x=0 this becomes 1=0; no multiplier exists



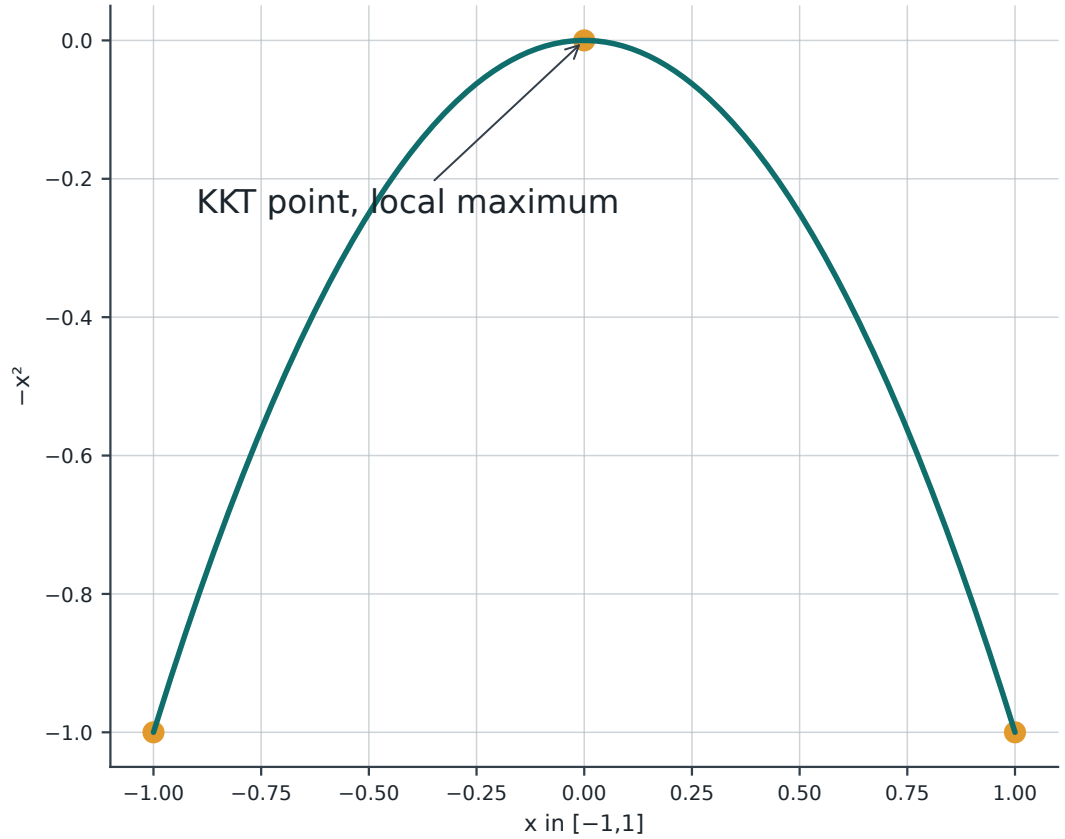
KKT does not certify a nonconvex minimum

$\min -x^2$ subject to $-1 \le x \le 1$

At x=0 both bounds are inactive, and λ=0 satisfies KKT

Yet x=0 is a local maximum; minima occur at x=±1

Check curvature and compare feasible values



Practice before the break

$\min (x-4)^2$ subject to $x \leq 1$

Write g, L and all KKT conditions

Find x^*, λ^* and f^*

Explain why $f'(x^*)$ is nonzero

Practice solution

$g = x - 1$; $L = (x-4)^2 + \lambda(x-1)$

Stationarity: $2(x-4) + \lambda = 0$

At the active bound: $x^* = 1$, $\lambda^* = 6$, $f^* = 9$

Convexity makes the verified KKT point globally optimal

Break · 10 minutes

Keep the sign convention in your notes

$\min f,\ g \le 0,\ L = f + \lambda^{T} g,\ \lambda \ge 0$

Next: how numerical methods use local models

A feasible line search checks two things

Require objective decrease and trial-point feasibility

For affine constraints, a feasible segment is easy to check

For nonlinear constraints, tangent directions may leave the feasible set

Practical methods may use merit functions or restoration steps

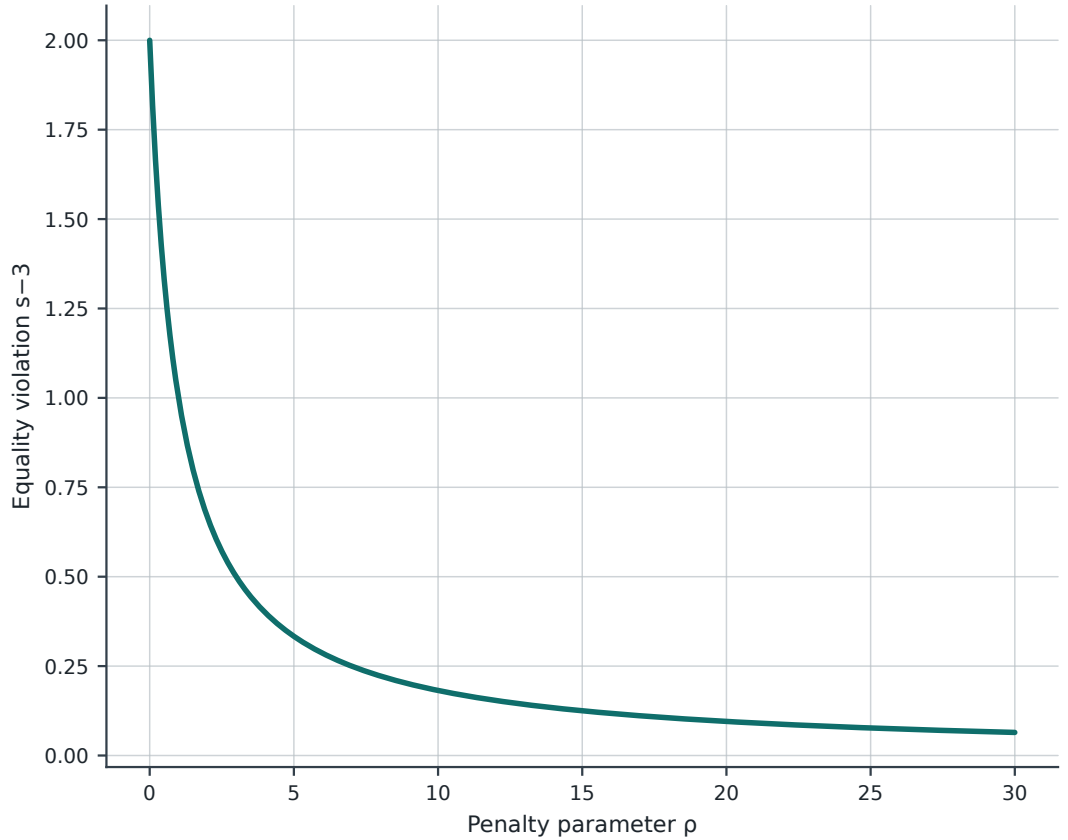
Quadratic penalty trades accuracy for conditioning

For $h(r)=0$, minimize $P_\rho = f(r) + (\rho/2)h(r)^2$

For our equality example,
$s = x + y = (5+3\rho)/(1+\rho)$

Residual $s - 3 = 2/(1+\rho)$, nonzero for finite ρ

Large ρ reduces violation but can worsen conditioning



Augmented Lagrangian combines two mechanisms

For equality $h=0$: $L_\rho = f + \nu h + (\rho/2)h^2$

Approximately minimize over r for a fixed ν

Update $\nu \leftarrow \nu + \rho h(r)$, then repeat

This is a method outline; stopping still needs residual checks

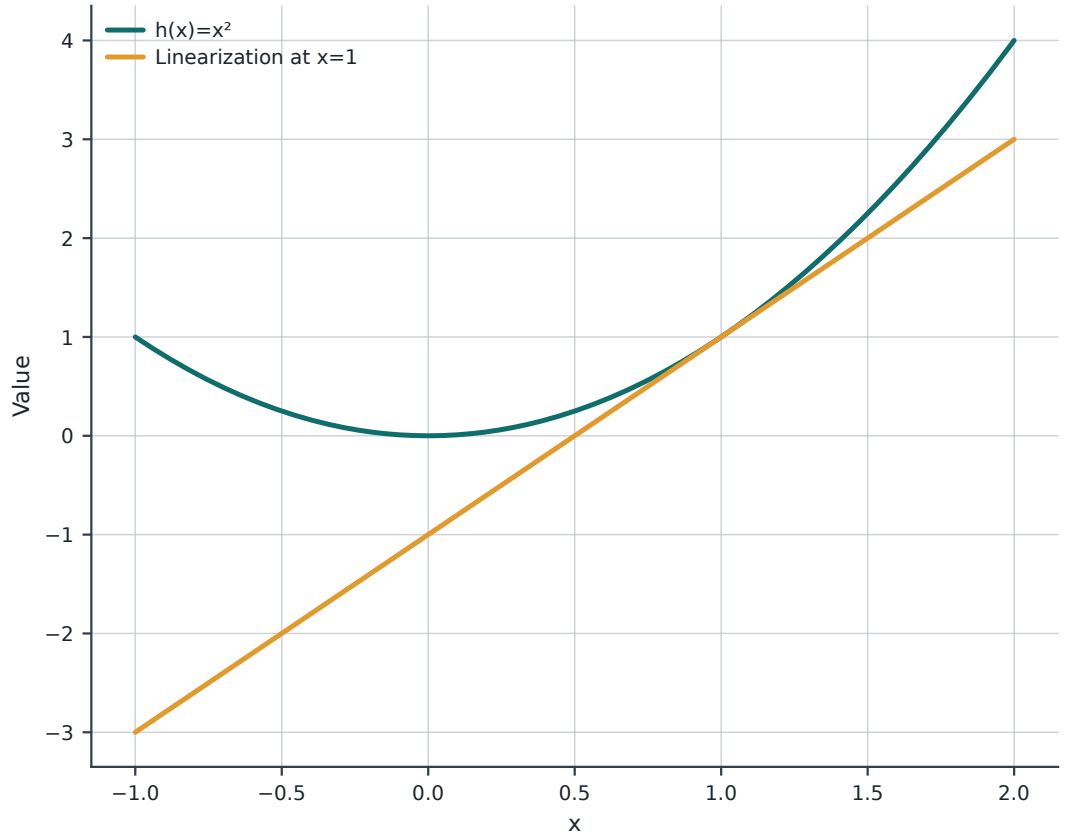
SQP solves a local constrained subproblem

Linearize constraints around the current point

Use a quadratic approximation of the Lagrangian

Solve for a trial direction, then control the step

Local approximations need repeated updates



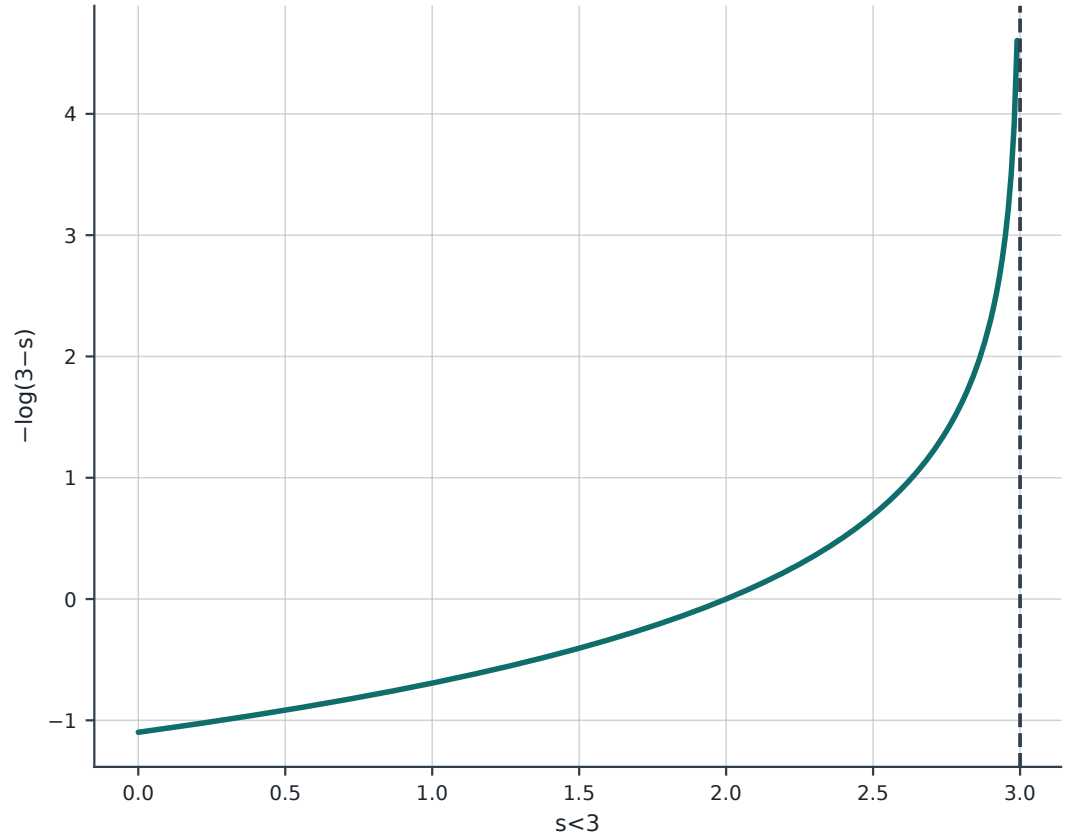
Interior-point methods approach from the interior

For $g_i<0$, a log barrier includes $-\mu\Sigma\log(-g_i)$

The barrier is undefined when a trial violates its domain

As μ decreases, solutions can approach an active boundary

A small μ alone does not certify convergence



Python uses a different inequality interface

Our mathematics: $g = x + y - 3 \leq 0$

SciPy SLSQP constraint dictionary: $c = 3 - x - y \geq 0$

Use bounds [(0,None),(0,None)]

Check the model convention before interpreting solver output

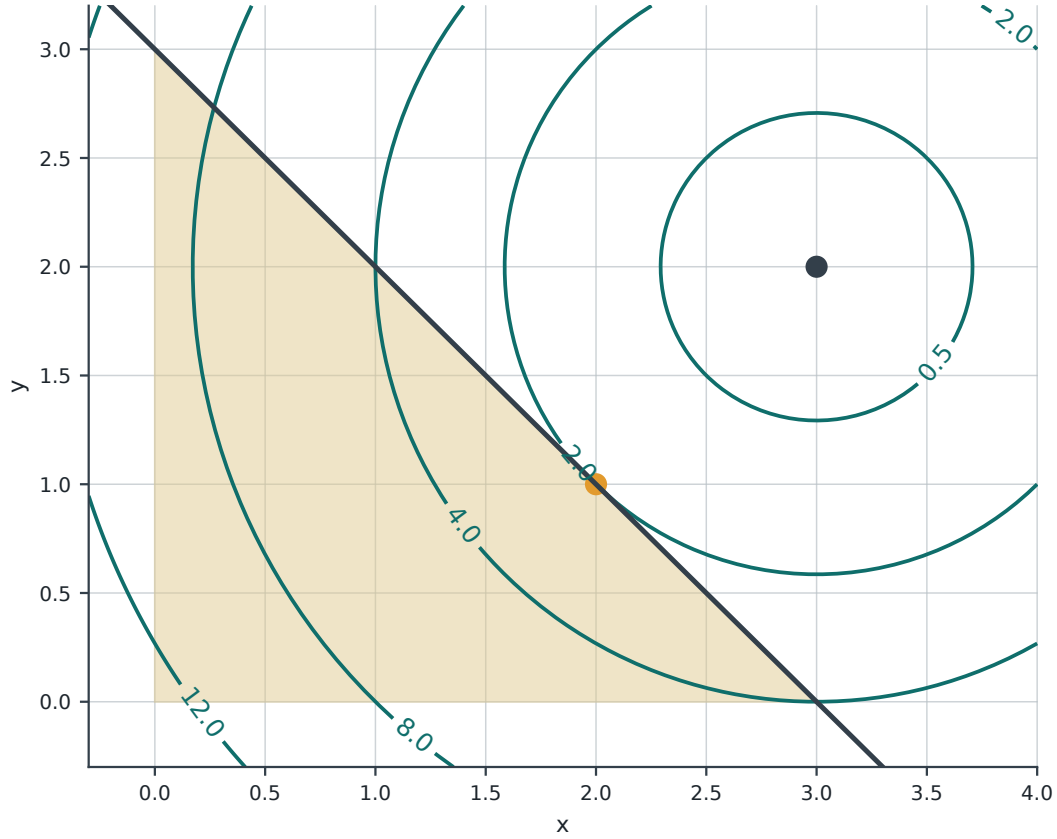
Verify a solver result independently

Read success and message, then recompute f and constraints

At the reference solution, infer $\lambda_1=2$ from stationarity

Report primal violation, stationarity and complementarity residuals

Repeat from several starts; compare with the analytic solution



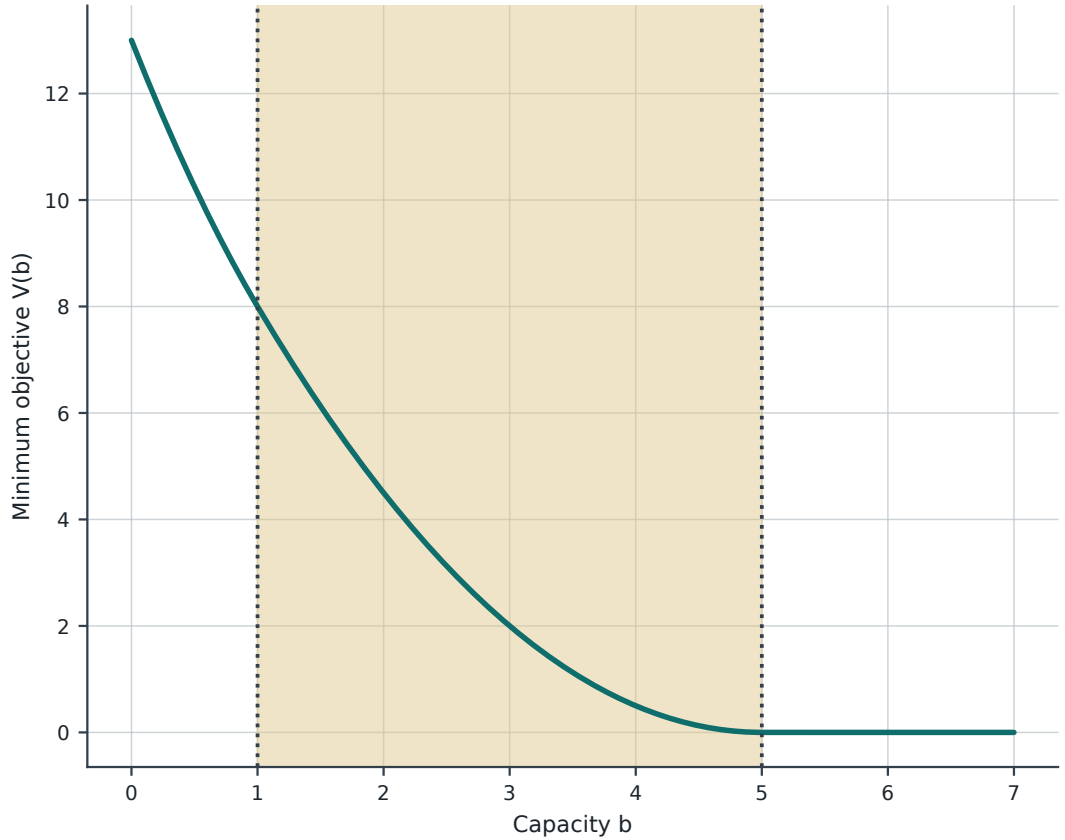
Capacity sensitivity changes at an active-set switch

For $1 \le b \le 5$: $x=(b+1)/2$, $y=(b-1)/2$

$V(b)=(b-5)^2/2$ and $\lambda=5-b$

For $b>5$: $(x,y)=(3,2)$, $V=0$ and $\lambda=0$

For $0 \le b<1$: $y=0$ becomes active



Derive the lower-capacity segment

At $0 \le b < 1$, set $y = 0$ and $x = b$

$V(b) = (b-3)^2 + 4$

$\lambda_1 = 6 - 2b$ and $\lambda_3 = 2 - 2b$; $\lambda_2 = 0$ is valid

For $0 < b < 1$, $dV/db = 2b - 6 = -\lambda_1$

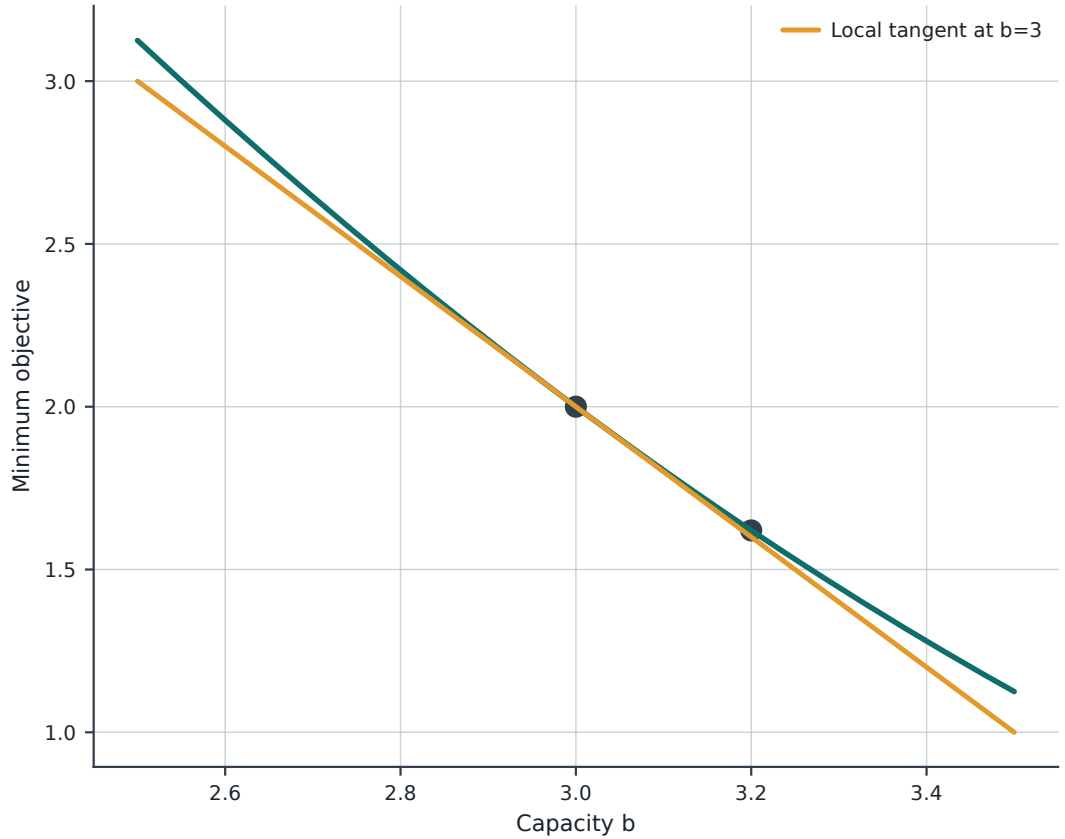
Worked sensitivity calculation

At b=3: x=2, y=1, V=2 and λ=2

At b=3.2: x=2.1, y=1.1 and V=1.62

First-order prediction: $\Delta V \approx -2(0.2) = -0.4$

Exact $\Delta V = -0.38$; nonlinear sensitivity is usually local



Choose a formulation before choosing a solver

LP: continuous variables with linear objective and constraints

MILP: linear expressions plus integer decisions

Convex NLP: verify convexity and model domains

Nonconvex NLP: local methods need careful interpretation

Integrated practice

A plant has continuous rates, a fixed start-up cost and nonlinear energy use

Which decisions need binary variables?

Can one local NLP run certify the mixed-integer global optimum?

Separate model construction, numerical solving and verification

What to show in the individual assignment

100 raw points across LP, MILP, unconstrained and constrained NLP

Submit derivations, reproducible code and a short interpretation

Show feasibility and residual checks alongside solver output

Deadline and course-grade weighting will be announced separately

Final examination preparation

90 minutes · 30 single-best-answer questions · 30 raw points

Scope: optimality, LP/duality, MILP, convexity and NLP methods

Include constrained NLP and KKT from this final class

Use a nonprogrammable calculator; no Python execution is needed

Final retrieval check

For $g \le 0$, what signs and zero products do KKT require?

Why can an optimal boundary point have $\nabla f \neq 0$?

Which evidence supports a global rather than local conclusion?

What calculation would you repeat to audit a solver answer?