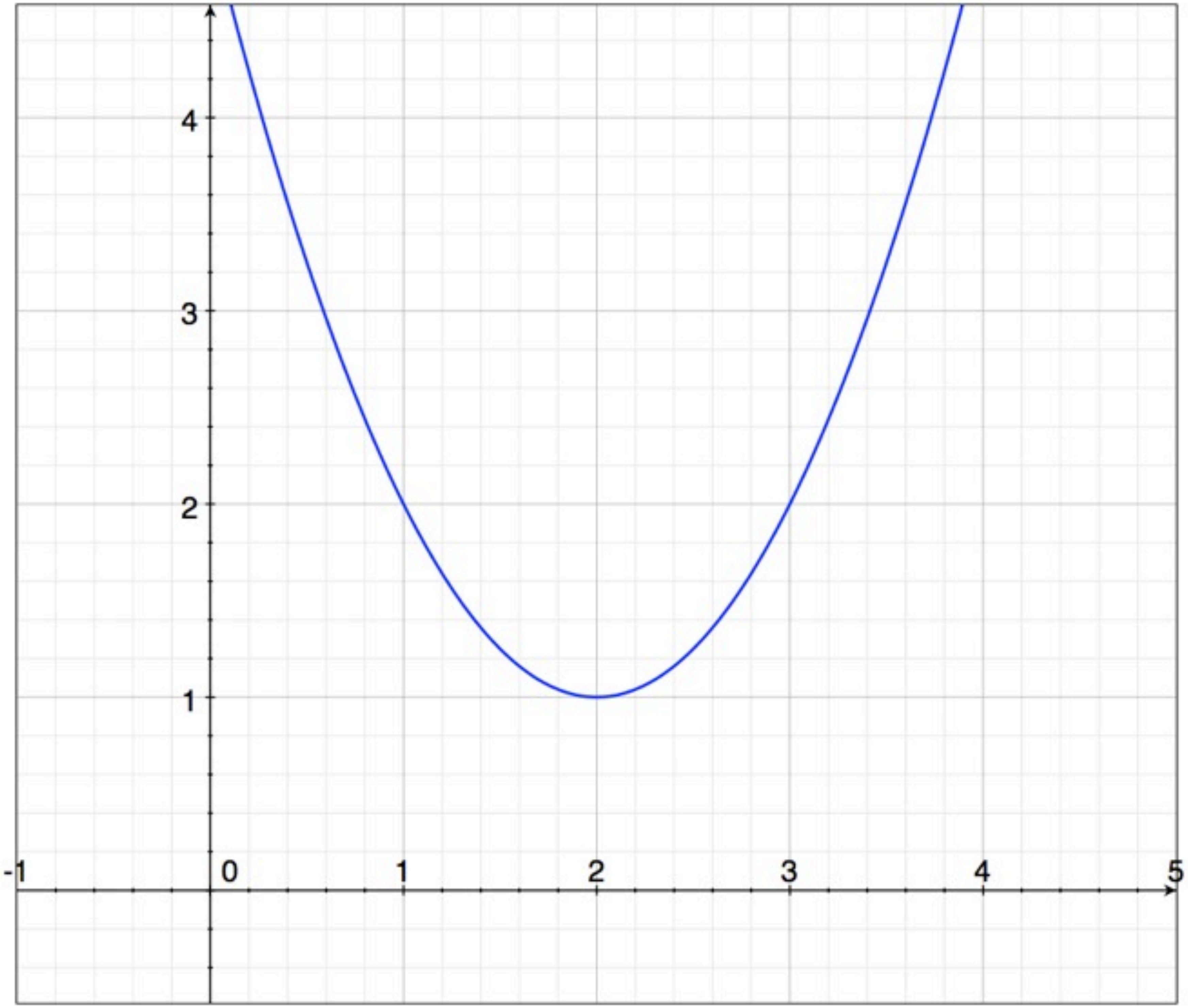
Constrained Optimization

- Optimize an objective function
- Subject to conditions expressed as equalities or inequalities

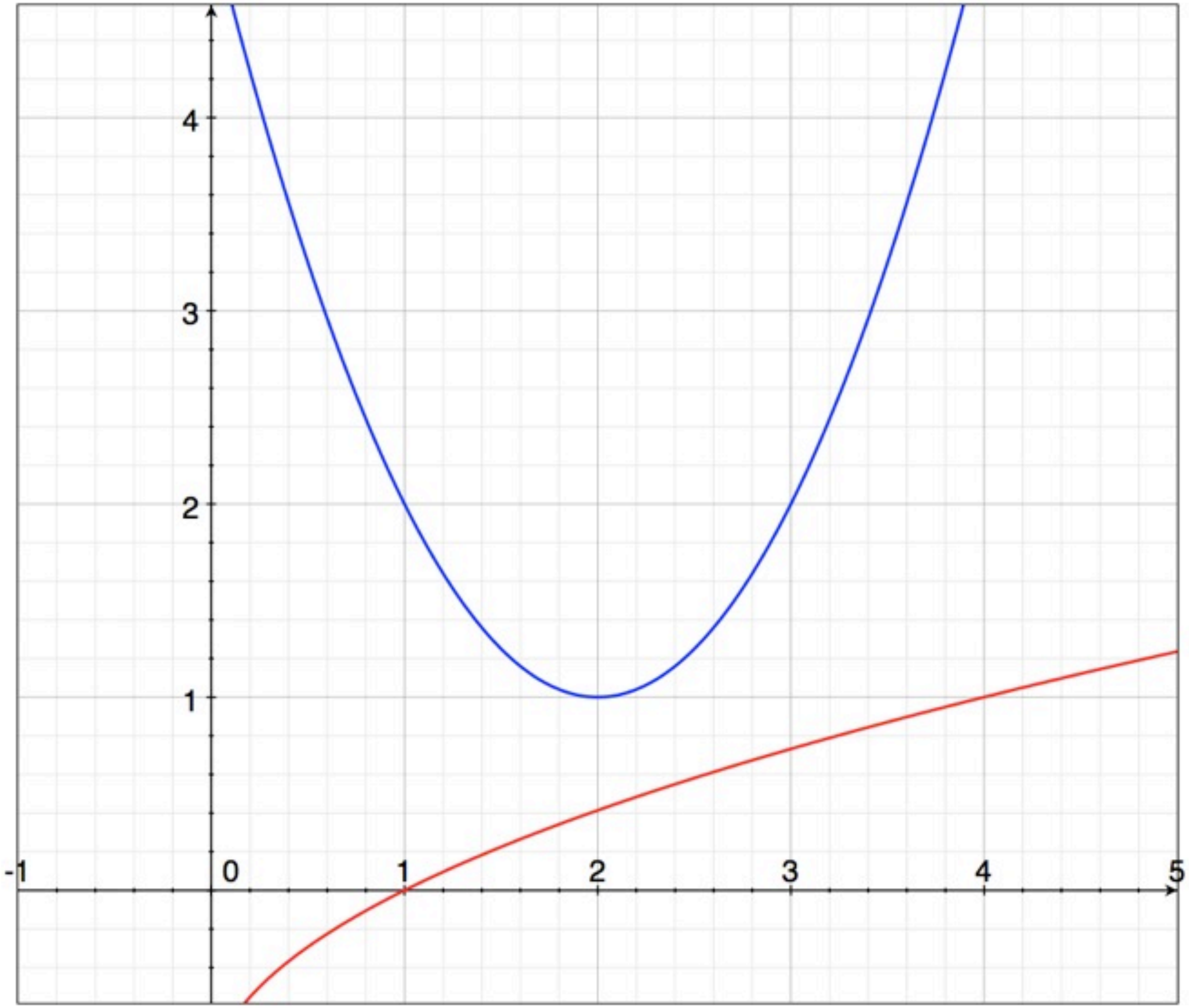
minimize	$f(x)$,	objective
w.r.t	x,	variables
subject to	$a \leq x \leq b$,	bound constraints
	$c_i(x) \leq u_i$,	inequality constraints
	$d_j(x) = v_j$.	equality constraints

Example: Univariate Constrained Optimization



Objective: $f(x) = (x - 2)^2 + 1$

Example: Univariate Constrained Optimization

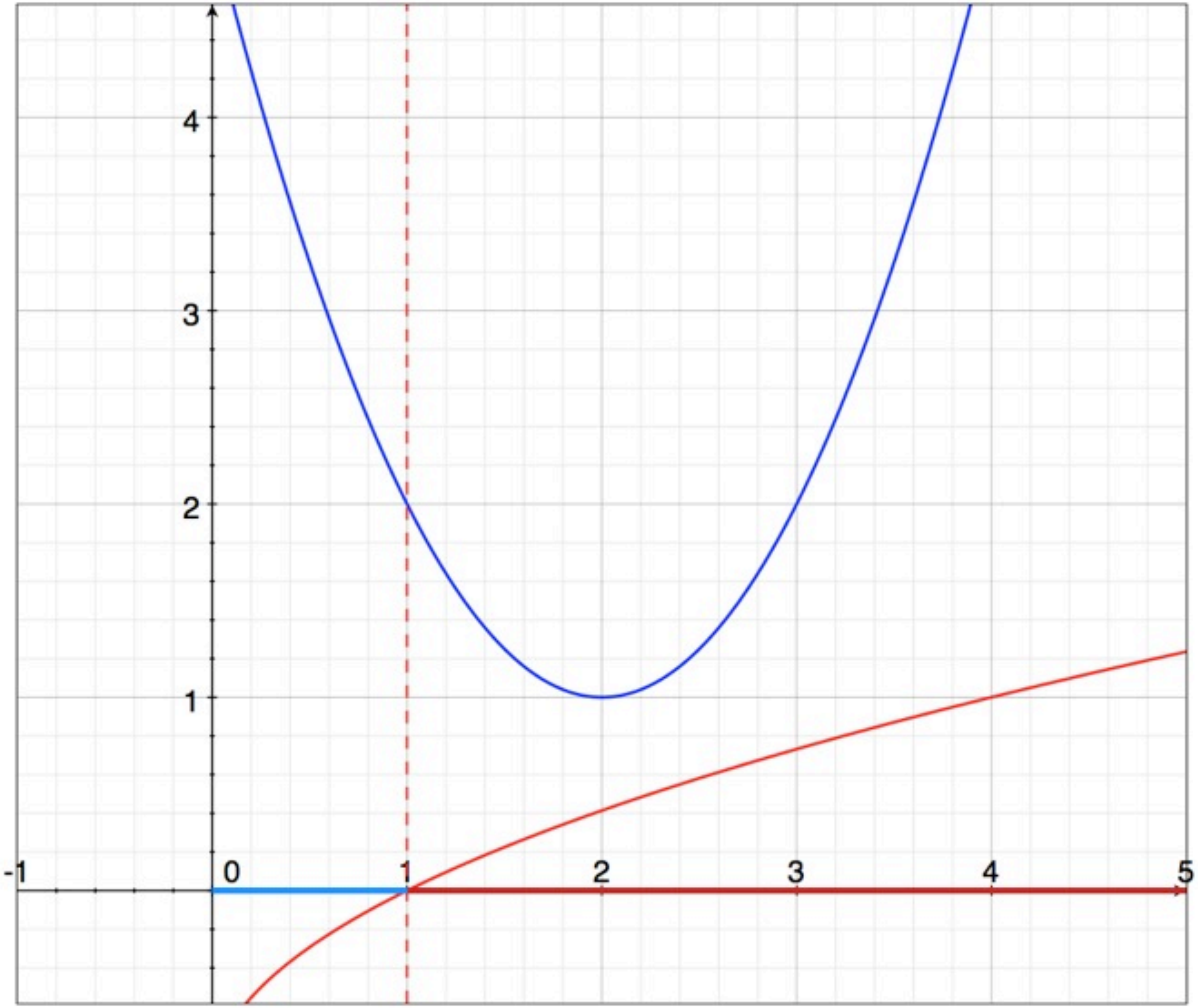


Objective: $f(x) = (x-2)^2 + 1$

Constraint: $c(x) = \sqrt{x} - 1 \leq 0$

Bound: $x \geq 0$

Example: Univariate Constrained Optimization



Objective: $f(x) = (x - 2)^2 + 1$

Constraint: $c(x) \leq \sqrt{x} - 1$

Bound: $x \geq 0$

Feasible Region: $0 \leq x \leq 1$

Infeasible Region: $x > 1$

Solution Methods

- Basic idea: convert to one or more unconstrained optimization problems
- Penalty function methods
  - Append a penalty for violating constraints (exterior penalty methods)
  - Append a penalty as you approach infeasibility (interior point methods)
- Method of Lagrange multipliers

Penalty-Function Methods

Objective

Penalty function

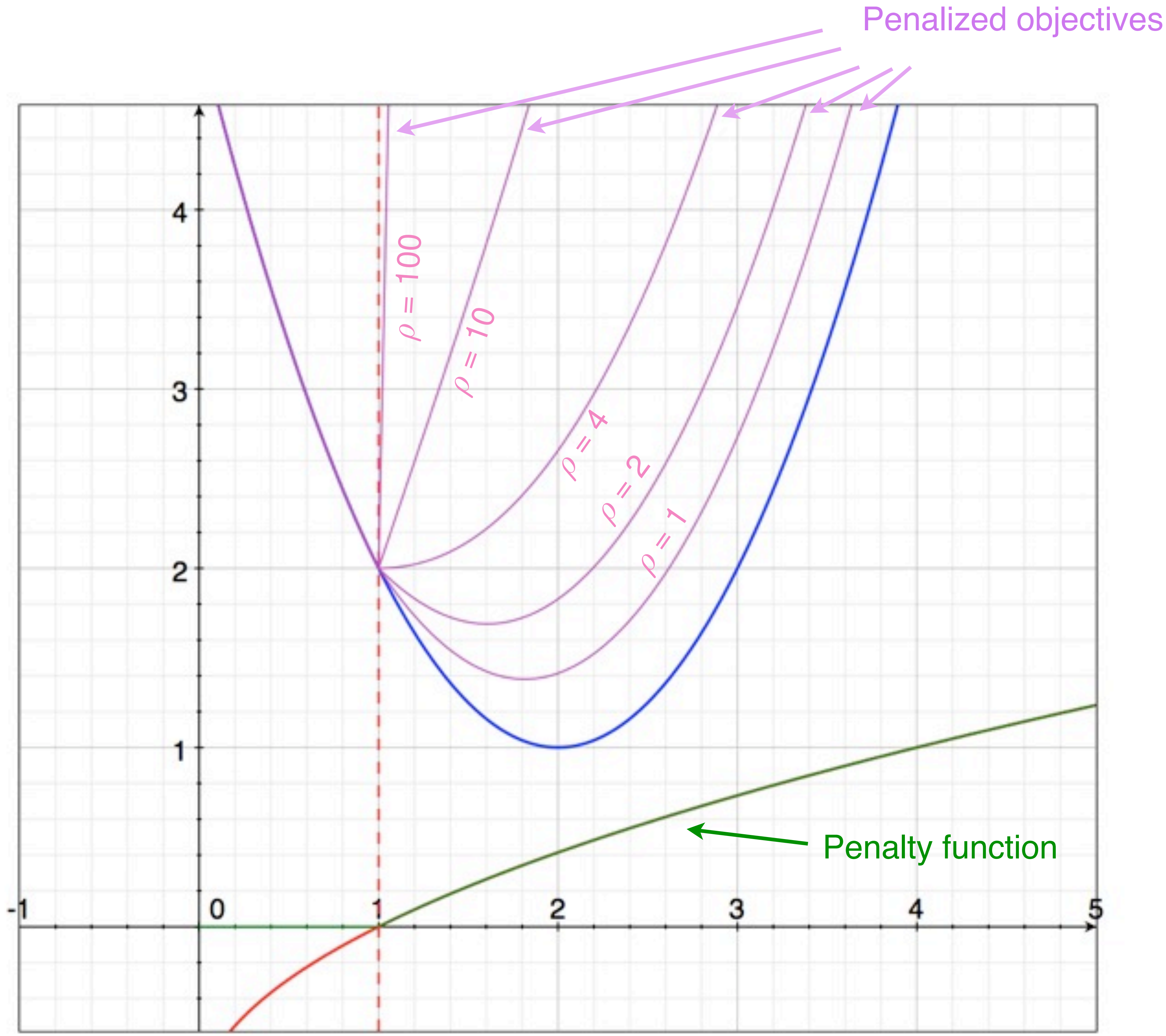
$$\pi(x, \rho) = \underbrace{f(x)}_{\text{Objective}} + \underbrace{\rho}_{\text{Penalty parameter (non-negative)}}\underbrace{\phi(x)}_{\text{Penalty function}}$$

Penalty parameter (non-negative)

1. Initialize penalty parameter
2. Initialize solution guess
3. Minimize penalized objective starting from guess
4. Update guess with the computed optimum
5. Go to 3., repeat

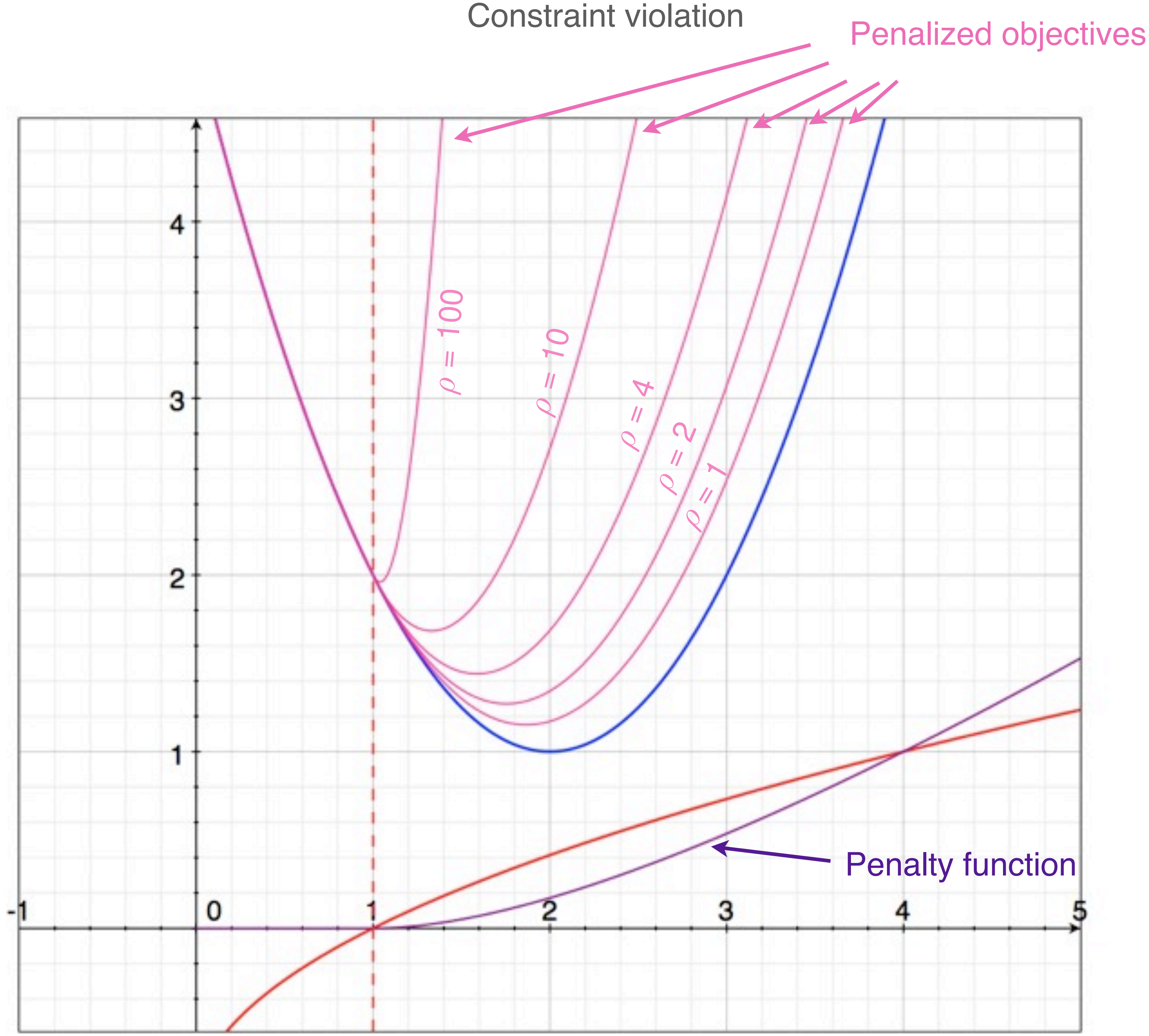
Linear Exterior Penalty Function

$$\phi_i(x) = \max(0, c_i(x) - u_i)$$



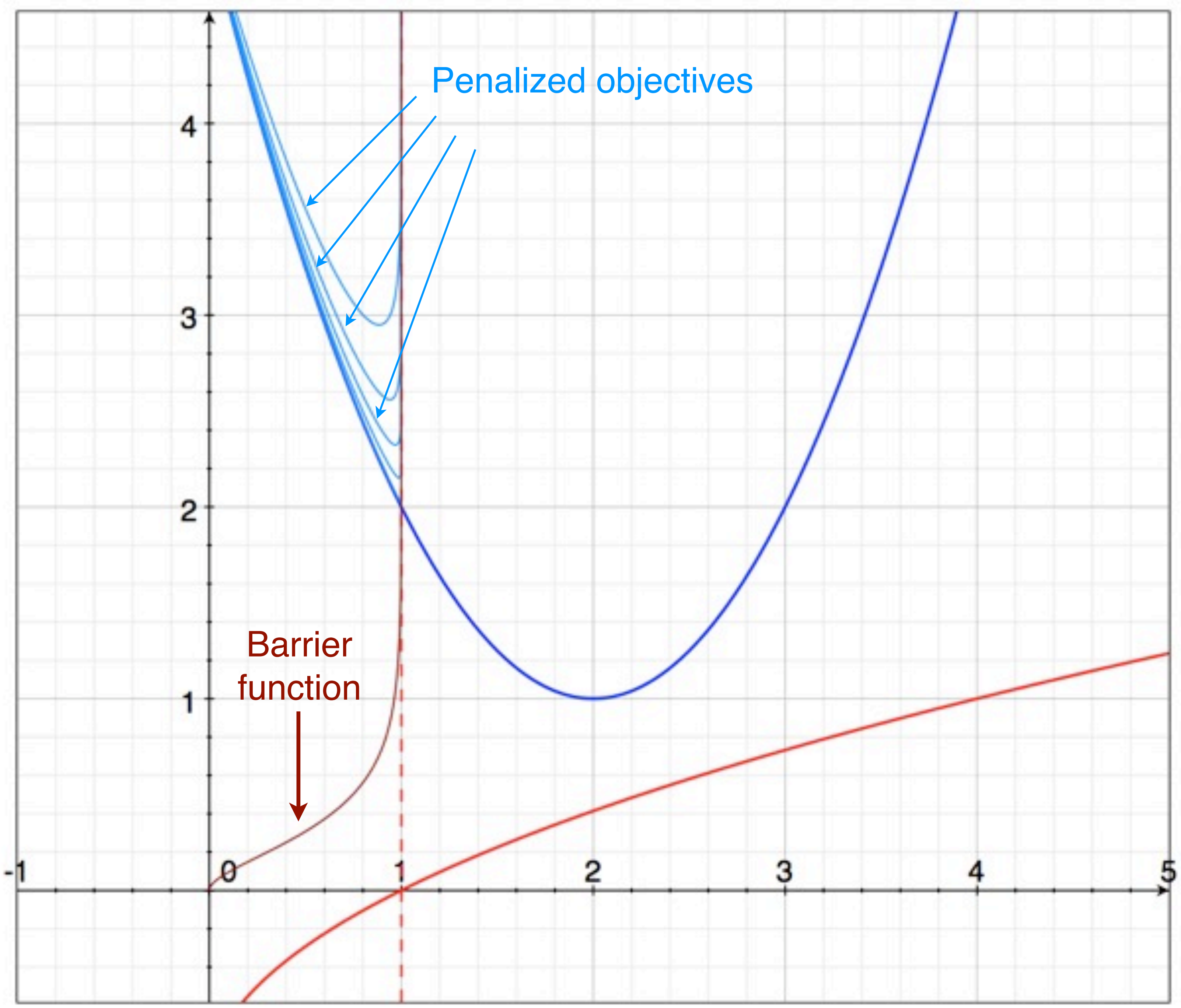
Quadratic Exterior Penalty Function

$$\phi_i(x) = \Big[\underbrace{\max(0, c_i(x) - u_i)}_{\text{Constraint violation}}\Big]^2$$



Interior-Point Methods

$$\pi(x, \mu) = f(x) - \underbrace{\mu}_{\text{Barrier parameter}} \overbrace{\log(u_i - c_i(x))}^{\text{Barrier function}}$$



Summary of Penalty Function Methods

- Quadratic penalty functions always yield slightly infeasible solutions
- Linear penalty functions yield non-differentiable penalized objectives
- Interior point methods never obtain exact solutions with active constraints
- Optimization performance tightly coupled to heuristics: choice of penalty parameters and update scheme for increasing them.
- Ill-conditioned Hessians resulting from large penalty parameters may cause numerical problems

Lagrange Multipliers: Introduction

- Powerful method with deep interpretations and implications
- Append each constraint function to the objective, multiplied by a scalar *for that constraint* called a Lagrange multiplier. This is the *Lagrangian* function

$$\mathcal{L}(x, \lambda) = f(x) + \sum_i \lambda_i (c_i(x) - u_i)$$

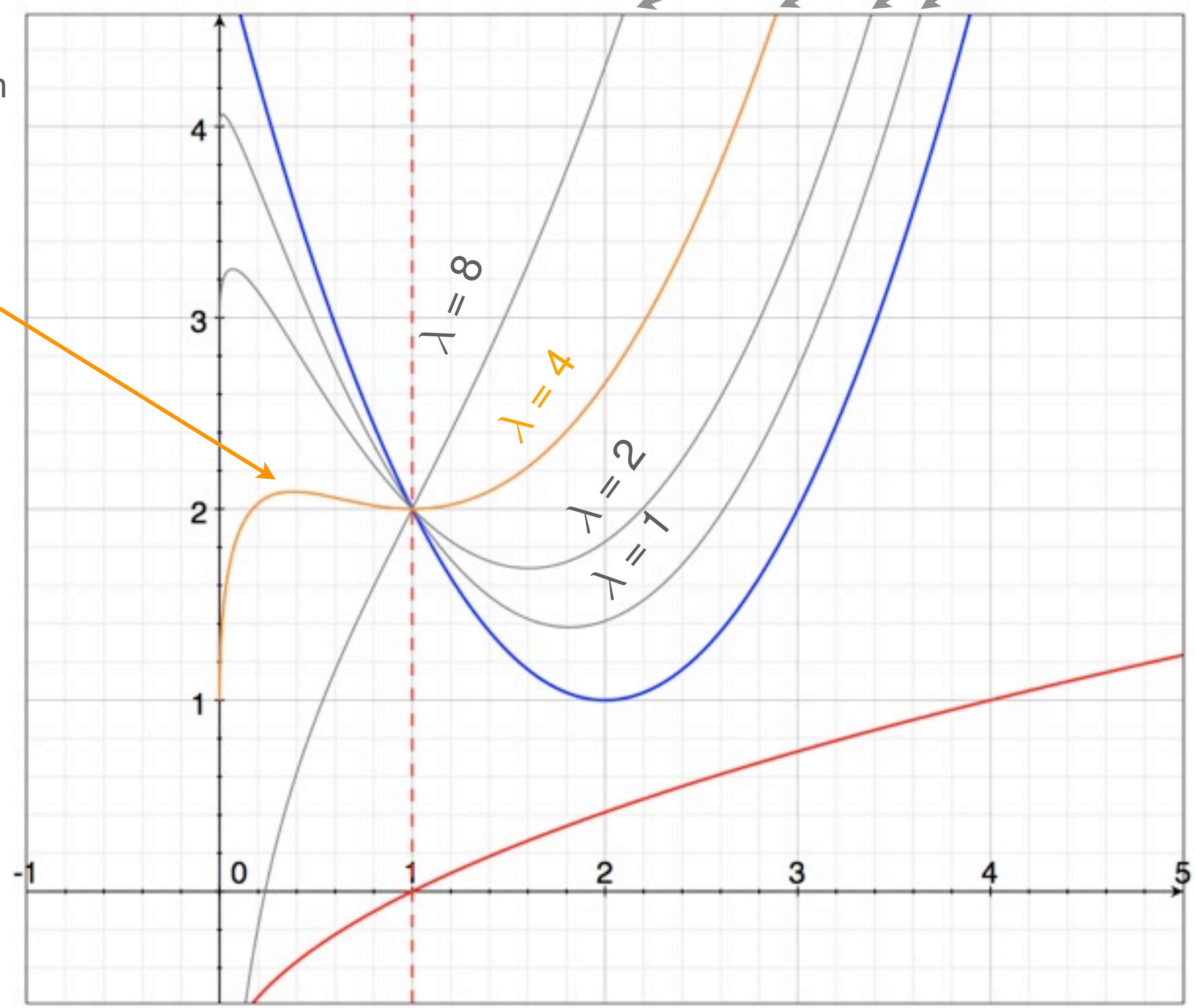
- Solution to the original constrained problem is deduced by solving for both an optimal *x* and an optimal set of Lagrange multipliers
- The original variables x are called primal variables, whereas the Lagrange multipliers are called dual variables
- Duality theory is both useful and beautiful, but beyond the scope of this class

Lagrange Multipliers: Motivation

$$\mathcal{L}(x, \lambda) = f(x) + \lambda c(x)$$

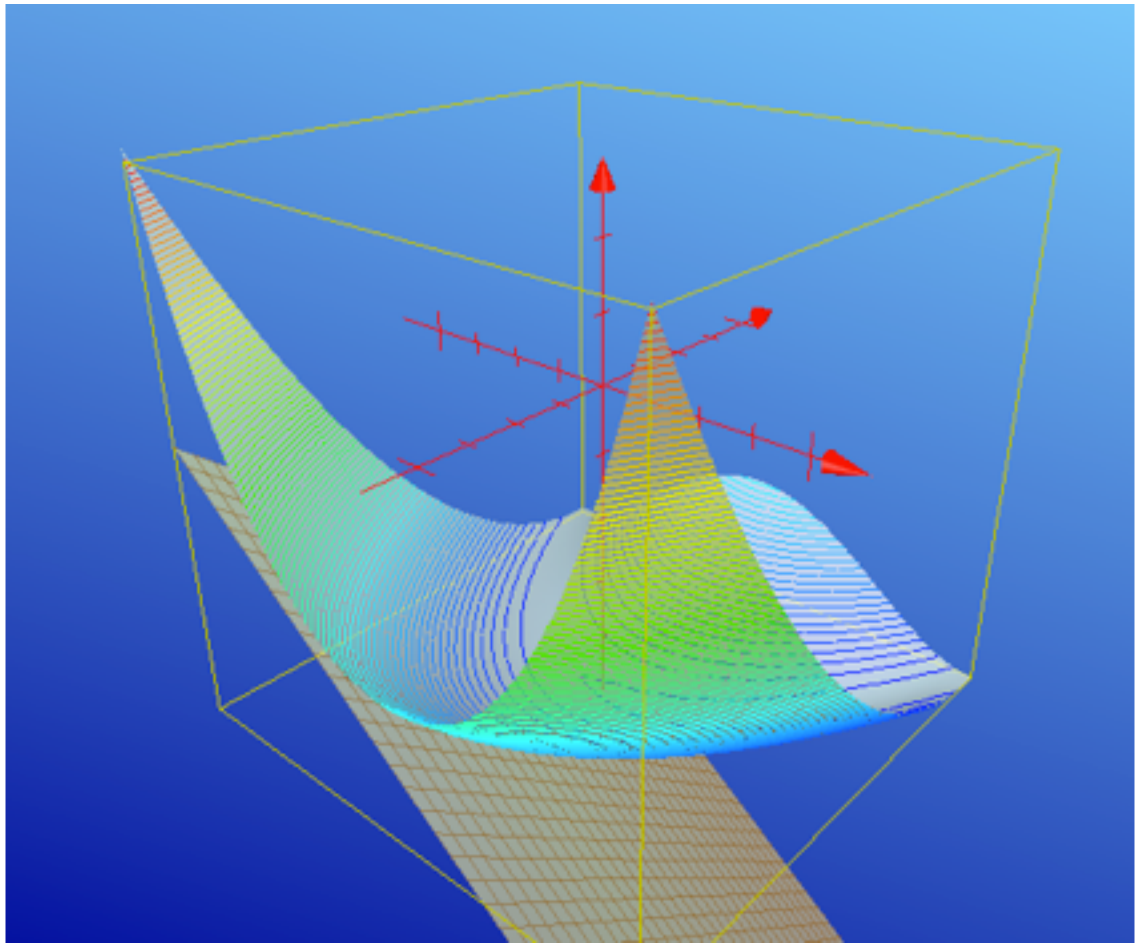
Lagrangians for various values of the Lagrange multiplier

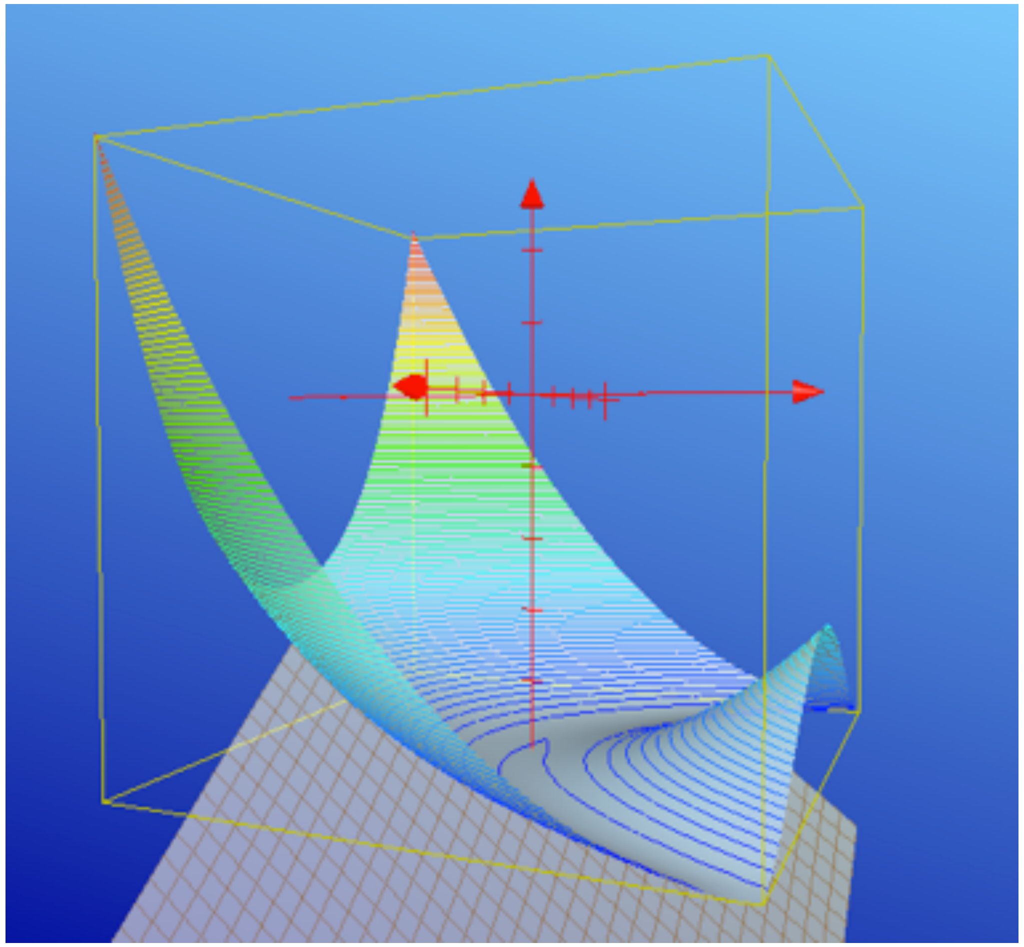
There is an optimal λ for which we obtain the constrained solution in x by minimizing the Lagrangian for that λ



On to Multivariate Problems

- What changes from univariate (1-D) to multivariate (n-D) problems?
- The little pictures you've been seeing get very complicated very quickly
- All the concepts still work, but need more careful treatment
- Absolutely essential are concepts of level curves and gradients





Level Curves and Gradients

- Consider a function $f : x \longrightarrow y$

- The level curves of *f* are curves in *x*-space

$$S_k = \{x \mid f(x) = k\}$$

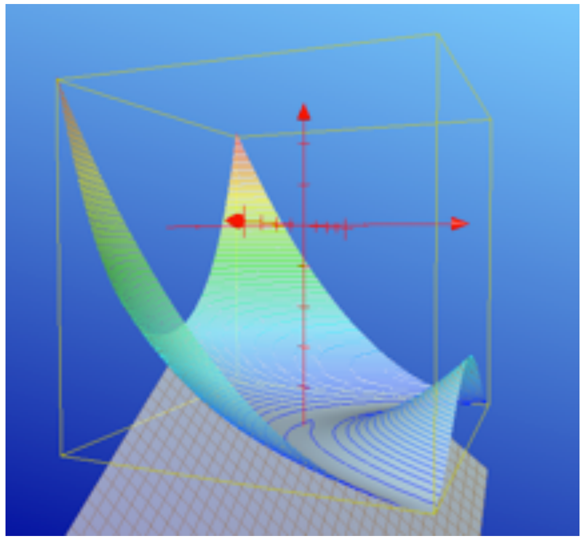
- The gradient of *f* w.r.t *x* is a vector in *x*-space

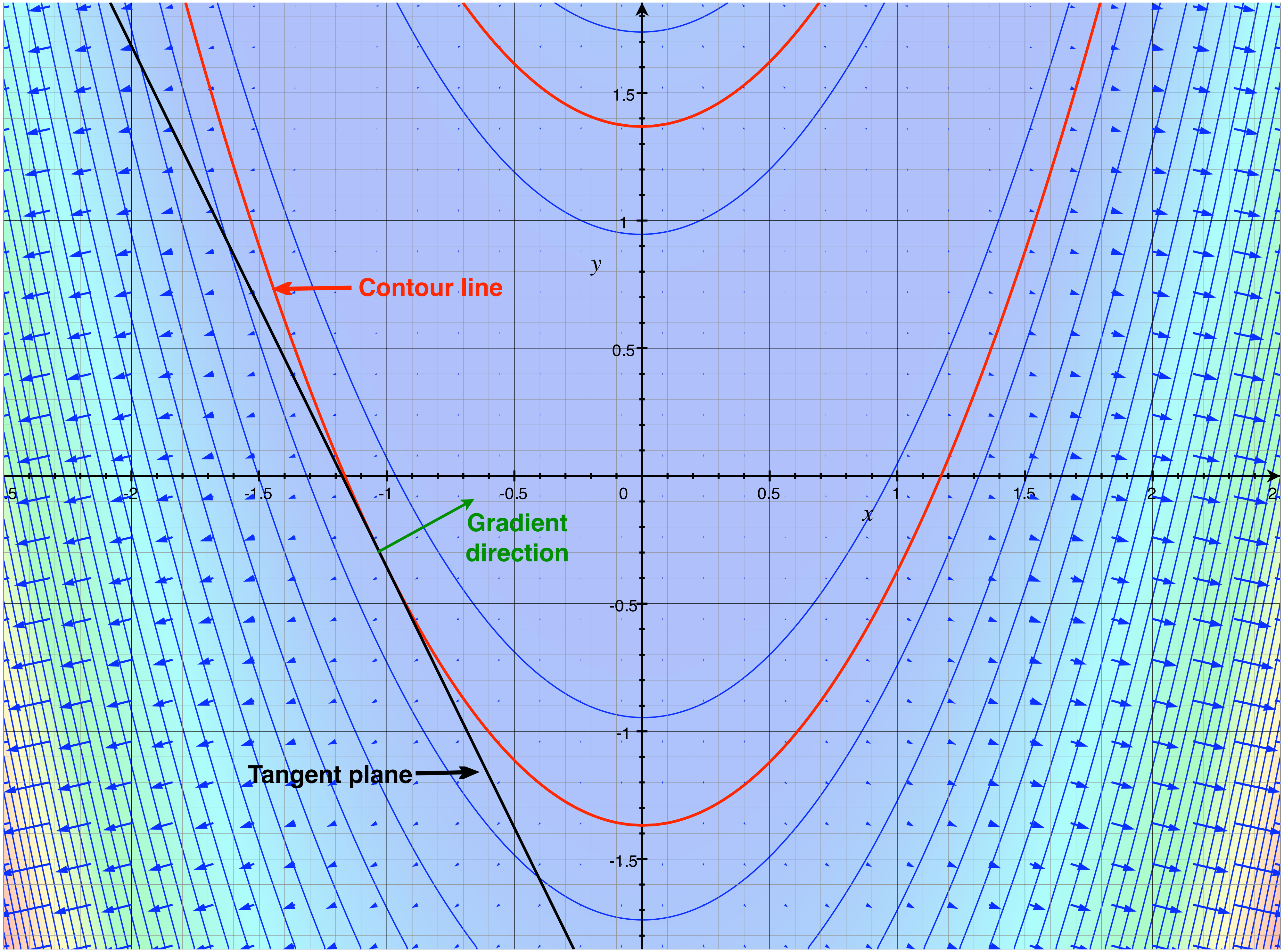
$$\nabla_x(f) = \left[\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right]$$

- Important fact from high-school calculus:

The gradient of a function is perpendicular to its level curves

Level Curves and Gradients





Gradients and First Order Changes

- Taylor series expansions: watch the dimensions of vectors and matrices!

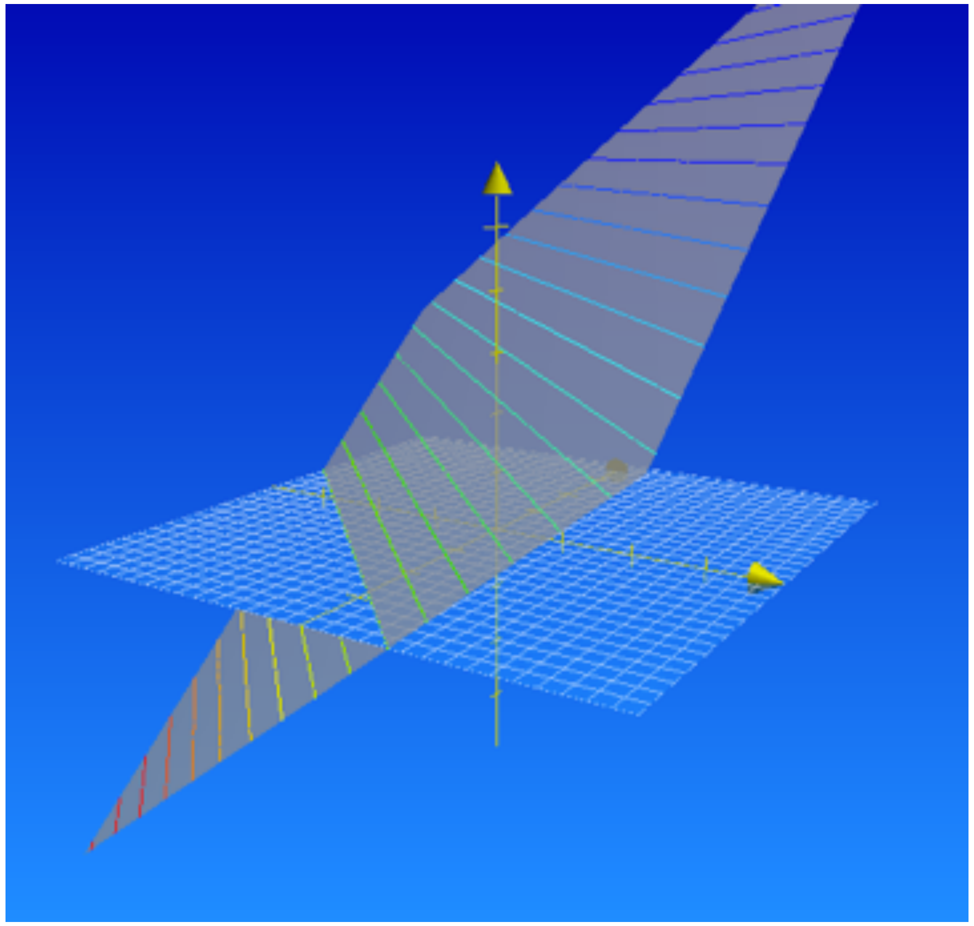
$$f(x_0 + \Delta x) = f(x_0) + \underbrace{\left[\nabla_x (f)\big|_{x_0} \right]^T}_{\text{gradient}} \Delta x + \frac{1}{2} \Delta x^T \underbrace{\nabla_x^2 (f)\Big|_{x_0}}_{\text{Hessian}} \Delta x + \mathcal{O}(\Delta x^3)$$

- The gradient defines local tangent plane and its ‘slope.’ We can deduce that
  - To first order, if a change in x has a component along the gradient, f will change
  - To first order, there is no change in f when moving perpendicular to the gradient
  - By definition, there is no change in f when moving along its level curve
  - Hence the level curve is perpendicular to the gradient
- The Hessian defines local curvature of f

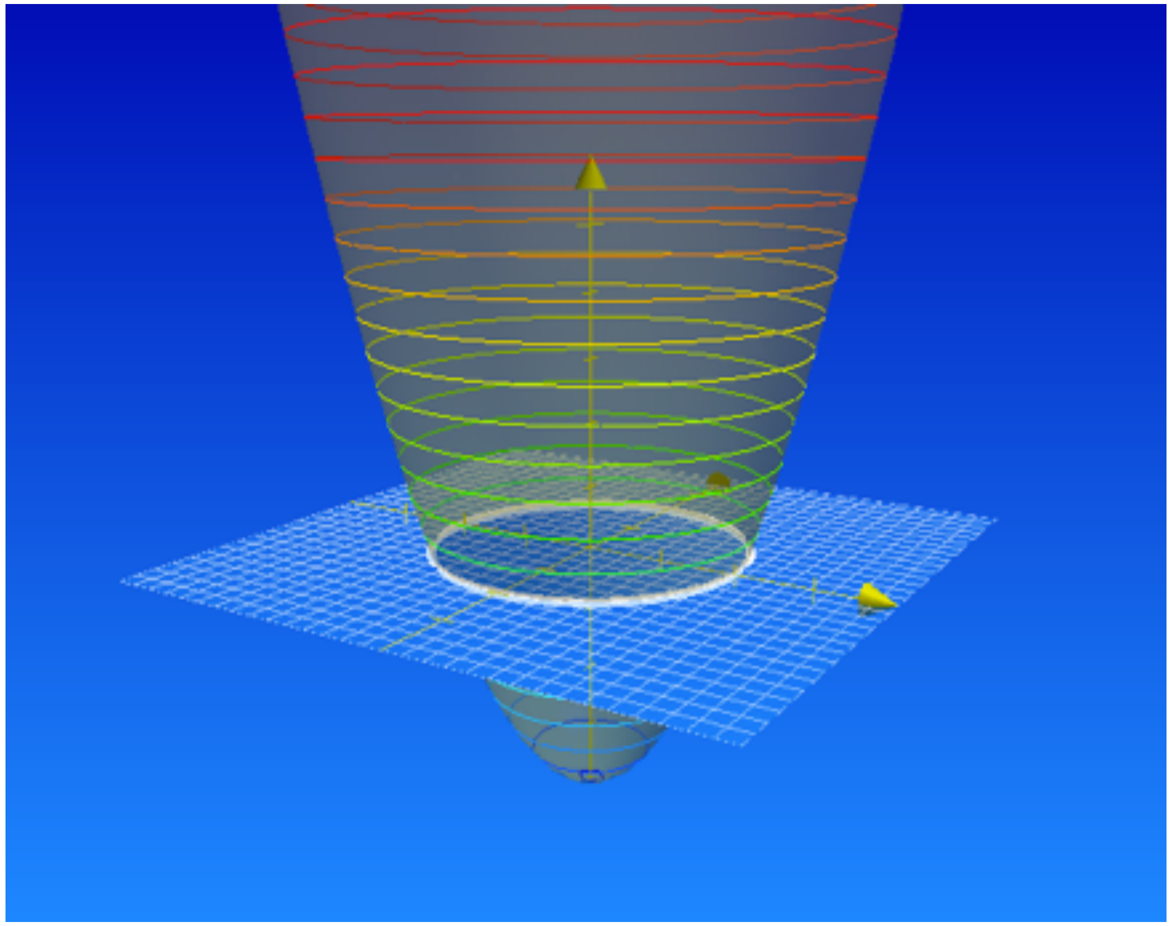
Multivariate Equality-Constrained Optimization

$$\begin{array}{ll} \text{minimize} & x_1 + x_2, \\ \text{subject to} & x_1^2 + x_2^2 = 4. \end{array}$$

Lowest point on this surface



Provided the value on this surface is 0

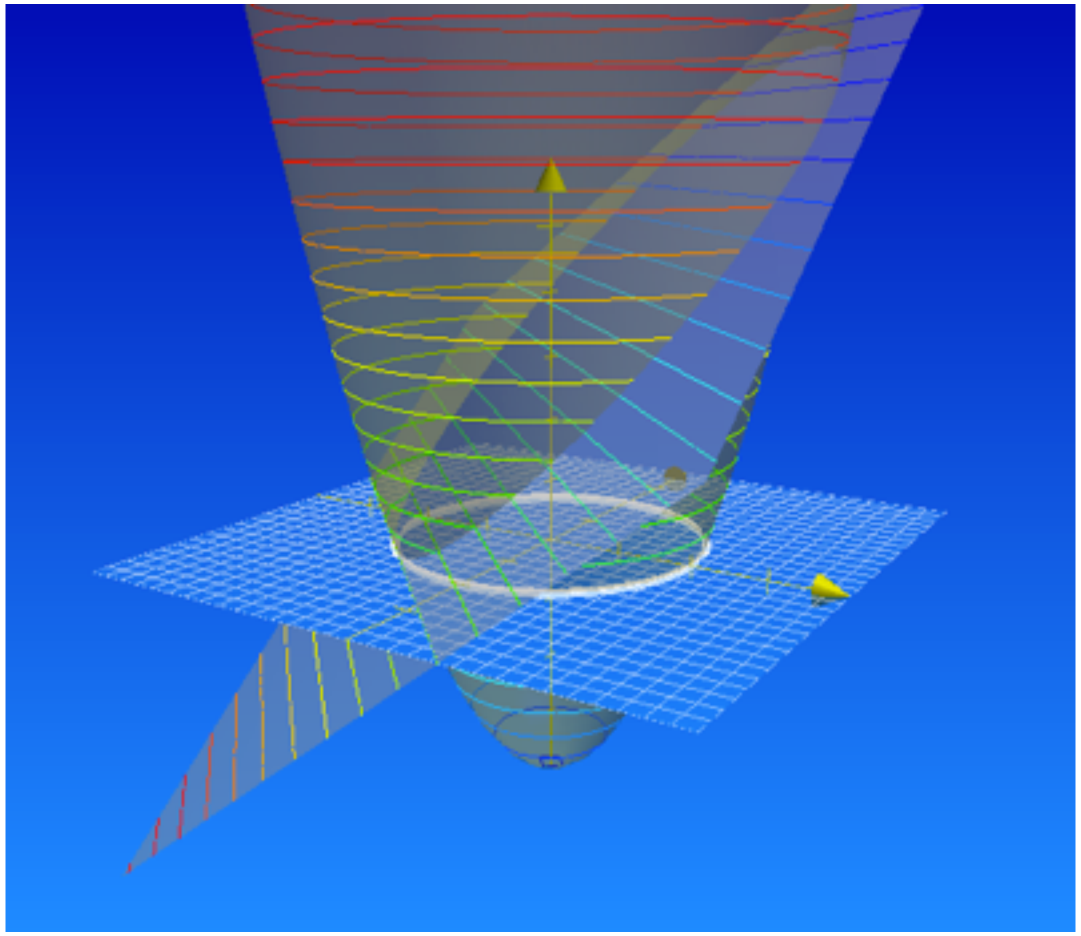


Conditions for a Constrained Extremum

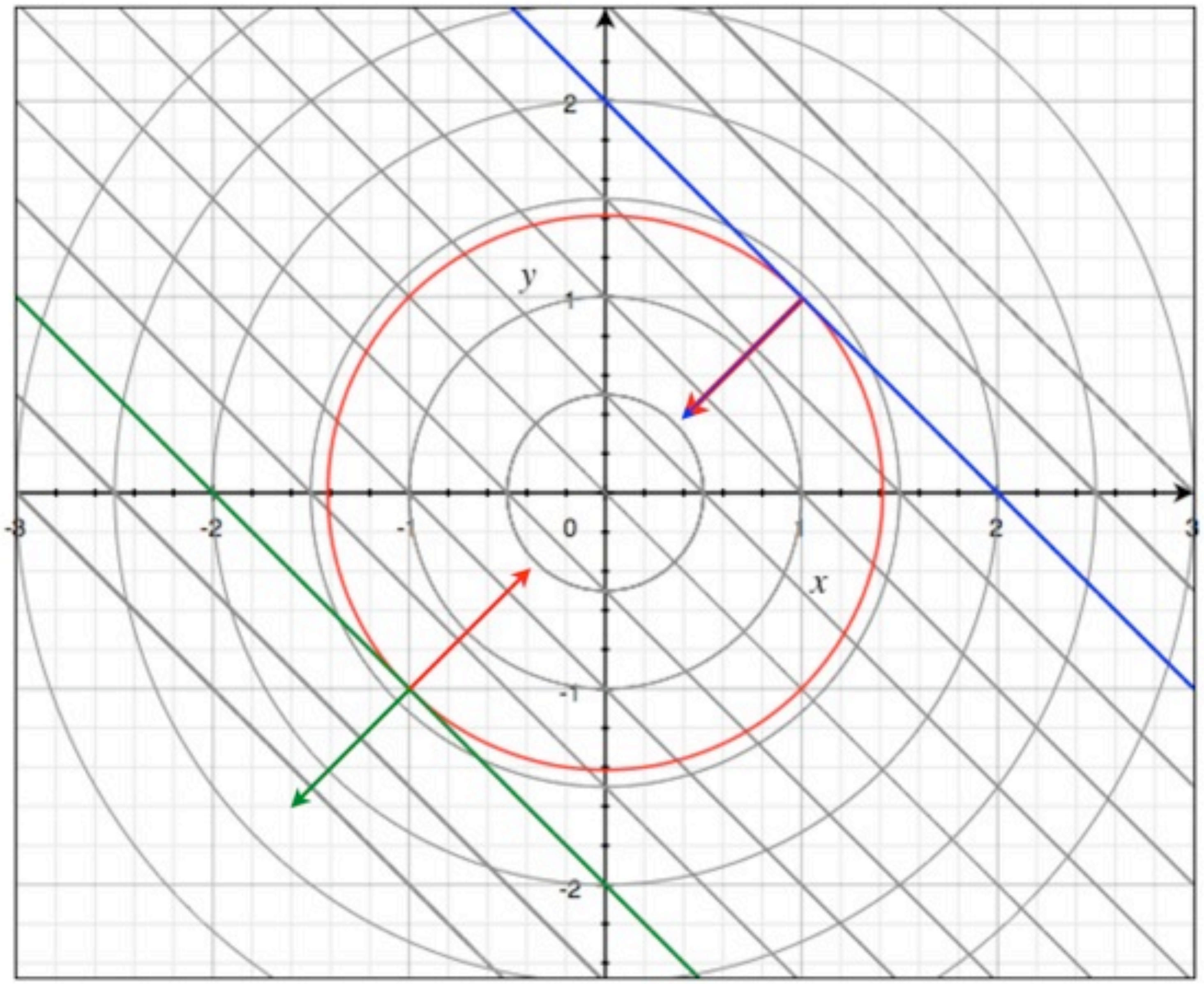
- Choose *x* anywhere on the circle, i.e., at a feasible point
- Any feasible small step in *x* must be perpendicular to the constraint gradient
- As long this step is not perpendicular to the objective gradient, we will get a change in *f*, and thereby, we at most have to reverse direction to reduce *f*
- The only way *f* can stop changing is when the step is perpendicular to both the objective and constraint gradients
- This means that the objective gradient and constraint gradient are parallel

$$\nabla_x(f)\big|_{x^\star} = \lambda \nabla_x(c)\big|_{x^\star}$$

- We have just found a constrained local extremum of *f*



Constrained Extrema



Question: what if some components of the constraint gradient are zero?