Constrained Optimization

Moritz Diehl

(some slide material was provided by W. Bangerth, K. Mombaur)

Nonlinear Programming (Problem Class 3)

- General problem formulation:

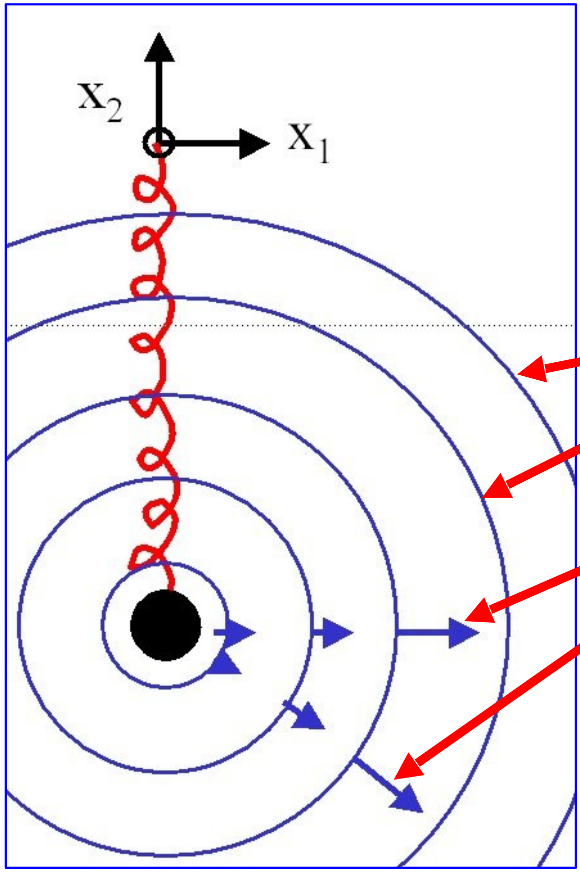
$$\boxed{\begin{array}{rclcl} \min f(x) & & & f: & D \subset R^n \rightarrow R \\ s.t.\, g(x) & = & 0 & g: & D \subset R^n \rightarrow R^l \\ h(x) & \geq & 0 & h: & D \subset R^n \rightarrow R^k \end{array}}$$

f objective function / cost function
g equality constraints
h inequality constraints

f,g,h shall be smooth (twice differentiable) functions

Recall: ball on a spring without constraints

$$\boxed{\min_{x \in R^2} x_1^2 + x_2^2 + m x_2}$$



contour lines of *f(x)*

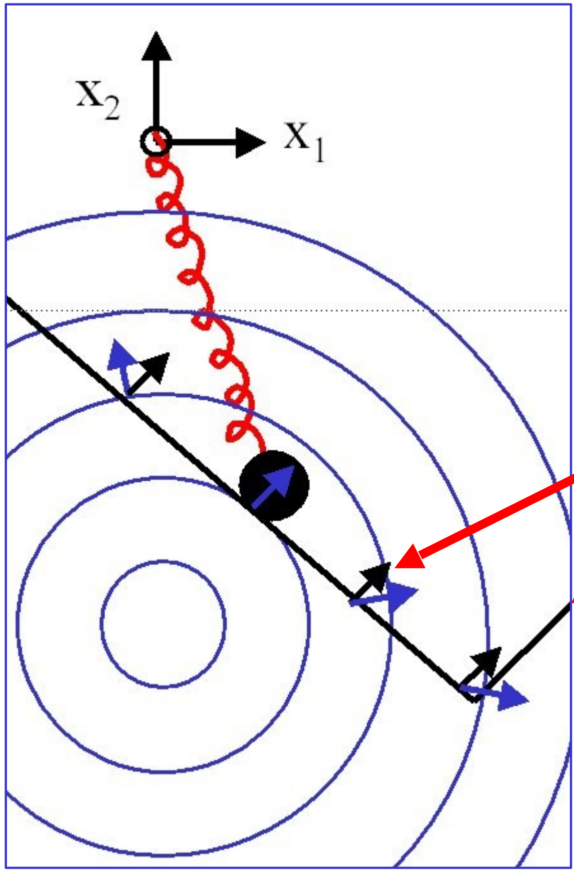
gradient vector

$$\nabla f(x) = (2x_1, 2x_2 + m)$$

unconstrained minimum:

$$0 = \nabla f(x^*) \Leftrightarrow (x_1^*, x_2^*) = (0, -\frac{m}{2})$$

Now: ball on a spring with constraints



$$\min f(x)$$
$$h_1(x) := 1 + x_1 + x_2 \quad \geq \quad 0$$
$$h_2(x) := 3 - x_1 + x_2 \quad \geq \quad 0$$

gradient ∇h_1 of active constraint

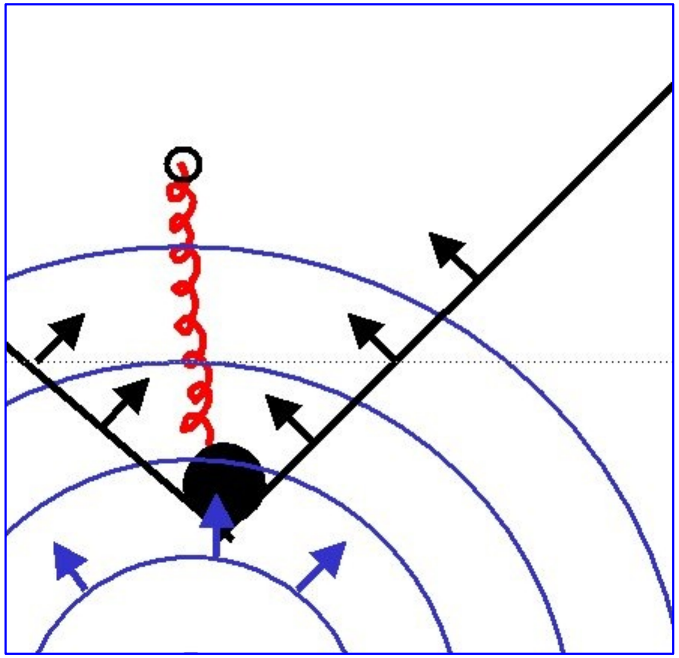
inactive constraint h_2

constrained minimum:

$$\boxed{\nabla f(x^*) = \mu_1 \nabla h_1(x^*)}$$

Lagrange multiplier

Ball on a spring with two active constraints



$$\min f(x)$$

$$h_1(x) := 1 + x_1 + x_2 \quad \geq \quad 0$$

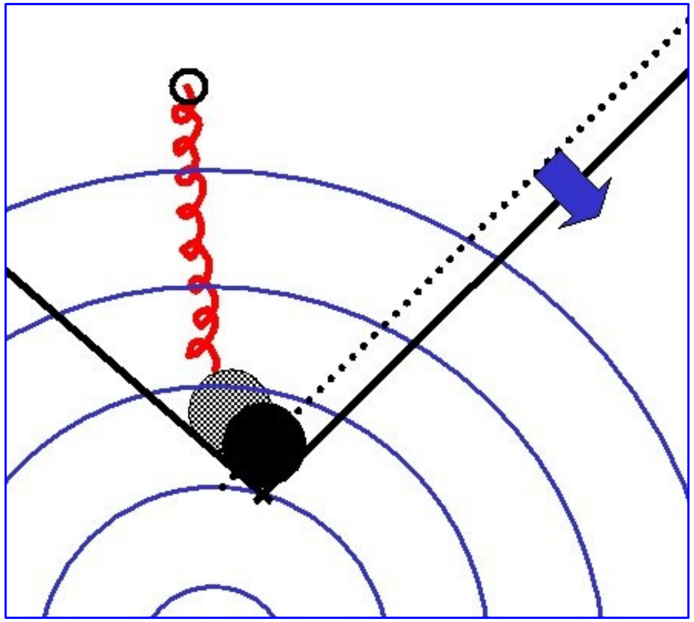
$$h_2(x) := 3 - x_1 + x_2 \quad \geq \quad 0$$

„equilibrium of forces“

$$\nabla f(x^*) = \mu_1 \nabla h_1(x^*) + \mu_2 \nabla h_2(x^*) \qquad \mu_1, \mu_2 \geq 0$$

„constraint forces“

Multipliers as „shadow prices“



old constraint: $h(x) \geq 0$

new constraint: $h(x) + \varepsilon \geq 0$

What happens if we relax a constraint?
Feasible set becomes bigger,
so new minimum $f(x_\varepsilon^*)$ becomes smaller.

How much would we gain?

$$\boxed{f(x_\varepsilon^*) \approx f(x^*) - \mu\varepsilon}$$

Multipliers show the hidden cost of constraints.

The Lagrangian Function

For constrained problems, introduce modification of objective function:

$$\boxed{L(x,\lambda,\mu) := f(x^*) - \sum \lambda_i g_i(x) - \sum \mu_i h_i(x)}$$

- equality multipliers λ_i may have both signs in a solution
- inequality multipliers μ_i cannot be negative (cf. shadow prices)
- for inactive constraints, multipliers μ_i are zero

Optimality conditions (constrained)

Karush-Kuhn-Tucker necessary conditions (KKT-conditions):

- x^* feasible
- there exist λ^*, μ^* such that

$$\boxed{\nabla_x L(x^*, \lambda^*, \mu^*) = 0}$$

$$(\Leftrightarrow \text{"Equilibrium"}\ \nabla f = \sum \lambda_i \nabla g_i + \sum \mu_i \nabla h_i)$$

- $\mu^* \geq 0$ holds
- and it holds the complementarity condition

$$\boxed{\mu^{*T} h(x^*) = 0}$$

i.e. $\mu_i^* = 0$ or $h_i(x^*) = 0$ for each i

Sequential Quadratic Programming (SQP)

Constrained problem:

$$\boxed{\begin{array}{rcl} & \min f(x) & \\ g(x) & = & 0 \\ h(x) & \geq & 0 \end{array}}$$

SQP Idea: Consider successively quadratic approximations of the problem:

$$\boxed{\begin{array}{l} \min\limits_{\Delta x} \; (\nabla f^k)^T \Delta x + \frac{1}{2} \Delta x^T H^k \Delta x \\ g(x^k) + \nabla g(x^k)^T \Delta x = 0 \\ h(x^k) + \nabla h(x^k)^T \Delta x \geq 0 \end{array}}$$

SQP method

- if we use the exact hessian of the Lagrangian

$$H = \nabla^2 L(x, \lambda, \mu)$$

  this leads to a newton-method for the optimality conditions and feasibility.
- with update-formulas for H^k, we obtain quasi-Newton SQP-methods.
- if we use appropriate update-formulas, we can have superlinear convergence.
- global convergence can be achieved by using a stepsize strategy.

SQP algorithm

0. Start with k=0, start value x^0 and $H^0=I$
1. Compute $f(x^k)$, $g(x^k)$, $h(x^k)$, $\nabla f(x^k)$, $\nabla g(x^k)$, $\nabla h(x^k)$
2. If x^k feasible and

$$\left\|\nabla L(x, \lambda, \mu)\right\| < \varepsilon$$

   then *stop* ➨ convergence achieved
3. Solve quadratic problem and get Δx^k
4. Perform line search and get stepsize t^k
5. Iterate

$$x^{k+1} = x^k + t^k \Delta x^k$$

6. Update hessian
7. k=k+1, goto step 1

Summary

- Lagrangian function plays important role in constrained optimization
- Lagrange multipliers of inequalities have positive sign
- KKT conditions are necessary optimality conditions