Lecture 18: Sequential Quadratic Programming

- Recap: Newton's method for solving nonlinear equations
- Recap: Equality-constrained QPs

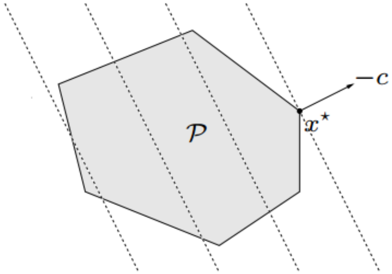
- SQP for *equality-constrained* nonlinear programming problems

Reference: N&W Ch.18-18.1

Types of constrained optimization problems

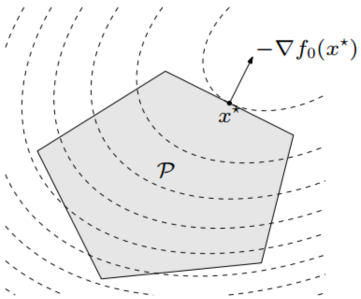
- Linear programming
  - Convex problem
  - Feasible set polyhedron

$$\begin{aligned} \text{minimize} \quad & c^\mathsf{T} x \\ \text{subject to} \quad & Ax \leq b \\ & Cx = d \end{aligned}$$



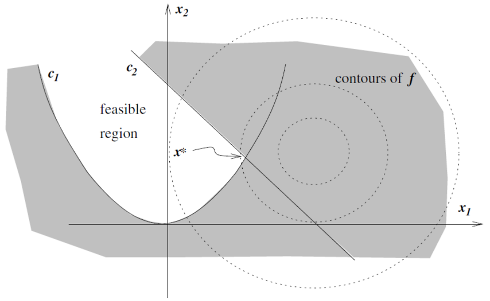
- Quadratic programming
  - Convex problem if $P \geq 0$
  - Feasible set polyhedron

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2} x^\mathsf{T} P x + q^\mathsf{T} x \\ \text{subject to} \quad & Ax \leq b \\ & Cx = d \end{aligned}$$



- Nonlinear programming
  - In general non-convex!

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g(x) = 0 \\ & h(x) \geq 0 \end{aligned}$$

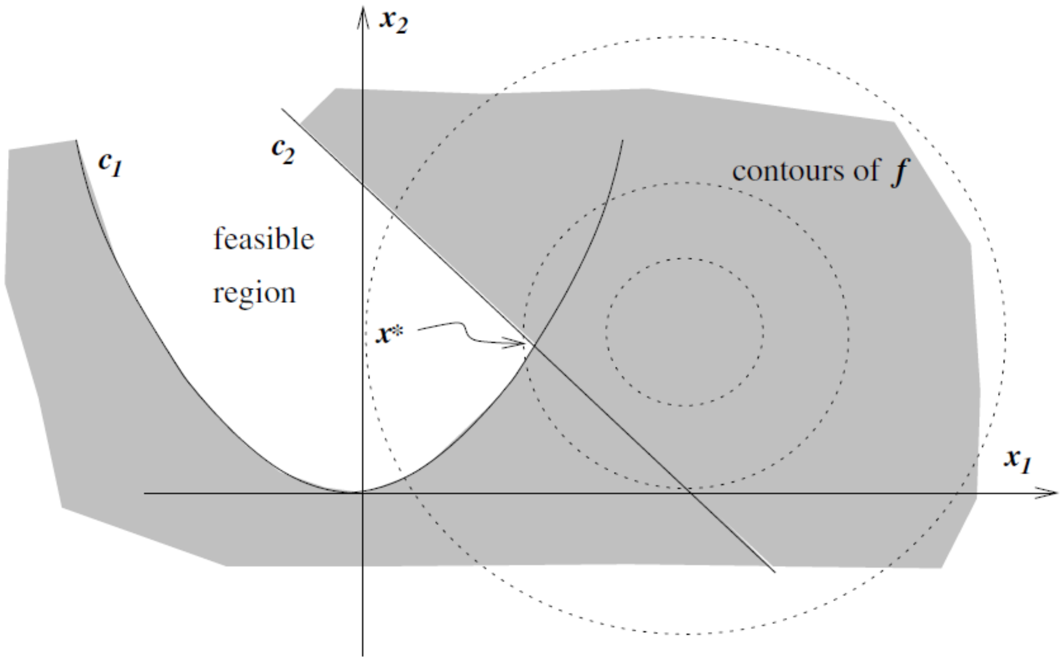


$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad \begin{cases} c_i(x) = 0, & i \in \mathcal{E}, \\ c_i(x) \geq 0, & i \in \mathcal{I}. \end{cases}$$

Nonlinear programming problem

$$\min_{x \in \mathbb{R}^n} f(x) \qquad \text{subject to} \quad \begin{aligned} c_i(x) &= 0, \quad i \in \mathcal{E}, \\ c_i(x) &\geq 0, \quad i \in \mathcal{I}. \end{aligned}$$

- Example: $\min\ (x_1 - 2)^2 + (x_2 - 1)^2 \qquad \text{subject to} \quad \begin{aligned} x_1^2 - x_2 &\leq 0, \\ x_1 + x_2 &\leq 2. \end{aligned}$



The Lagrangian

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad \begin{cases} c_i(x) = 0, & i \in \mathcal{E} \\ c_i(x) \geq 0, & i \in \mathcal{I} \end{cases}$$

For constrained functions, introduce modification of objective function (the *Lagrangian*):

$$\mathcal{L}(x, \lambda) = f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x)$$

- Multipliers for *equality* constrains may have both signs in a solution
- Multipliers for *inequality* constraints cannot be negative (cf. shadow prices)
- For (inequality) constraints that are *inactive*, multipliers are zero

KKT conditions (Theorem 12.1)

$$\min_{x\in\mathbb{R}^n} f(x) \quad \text{subject to} \quad \begin{cases} c_i(x) = 0, & i \in \mathcal{E} \\ c_i(x) \geq 0, & i \in \mathcal{I} \end{cases}$$

KKT-conditions (First-order necessary conditions): If x^* is a local solution and LICQ holds, then there exist λ^* such that

$$\begin{aligned} \nabla_x \mathcal{L}(x^*, \lambda^*) &= 0, && && \text{(stationarity)} \\ \left.\begin{aligned} c_i(x^*) &= 0, \quad \forall i \in \mathcal{E}, \\ c_i(x^*) &\geq 0, \quad \forall i \in \mathcal{I}, \end{aligned}\right\} && && && \text{(primal feasibility)} \\ \lambda_i^* &\geq 0, && \forall i \in \mathcal{I}, && \text{(dual feasibility)} \\ \lambda_i^* c_i(x^*) &= 0, && \forall i \in \mathcal{E} \cup \mathcal{I}. && \text{(complementarity condition/ complementary slackness)} \end{aligned}$$

Either $\lambda_i^* = 0$ *or* $c_i(x^*) = 0$

(*strict* complimentarity: Only one of them is zero)

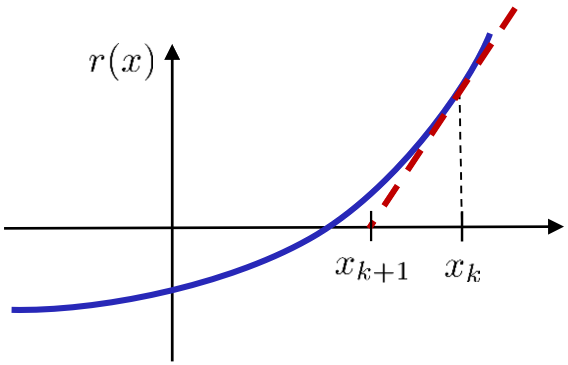
Newton's method for solving nonlinear equations (Ch. 11)

- Solve equation system $r(x) = 0, \quad r(x) : \mathbb{R}^n \to \mathbb{R}^n$
- Assume Jacobian $J(x) \in \mathbb{R}^{n \times n}$ exists and is continuous $J(x) = \begin{pmatrix} \frac{\partial r_1}{\partial x_1} & \frac{\partial r_1}{\partial x_2} & \cdots \\ \frac{\partial r_2}{\partial x_1} & \frac{\partial r_2}{\partial x_2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$
- Taylor: $r(x+p) = r(x) + J(x)p + O(\|p\|^2)$

Algorithm 11.1 (Newton's Method for Nonlinear Equations).
Choose x_0;
for $k = 0, 1, 2, \ldots$
  Calculate a solution p_k to the Newton equations

$$J(x_k)p_k = -r(x_k);$$

  $x_{k+1} \leftarrow x_k + p_k$;
end (for)



- (Local) convergence rate (Thm 11.2): Quadratic convergence if $J(x)$ is Lipschitz continuous (that is, very good convergence rate)

- If we set $r(x) = \nabla f(x)$, then this method corresponds to Newton's method for minimizing $f(x)$

$$p_k = -J(x_k)^{-1} r(x_k) \quad \longleftrightarrow \quad p_k = -\left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k)$$

Equality-constrained QP (EQP)

$$\min_{x\in\mathbb{R}^n} \quad \frac{1}{2}x^\top G x + c^\top x$$
$$\text{subject to} \quad Ax = b, \quad A \in \mathbb{R}^{m\times n}$$

Basic assumption:
A full row rank

- KKT-conditions (KKT system, KKT matrix):

$$\begin{pmatrix} G & -A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} x^* \\ \lambda^* \end{pmatrix} = \begin{pmatrix} -c \\ b \end{pmatrix} \quad \text{or, if we let } x^* = x + p, \quad \begin{pmatrix} G & A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} -p \\ \lambda^* \end{pmatrix} = \begin{pmatrix} c + Gx \\ Ax - b \end{pmatrix}$$

- Solvable when $Z^\top G Z > 0$ (columns of *Z* basis for nullspace of *A*):

$$Z^\top G Z > 0 \overset{\text{Lemma 16.1}}{\Rightarrow} K = \begin{pmatrix} G & A^\top \\ A & 0 \end{pmatrix} \text{ non-singular}$$
$$\Rightarrow \begin{pmatrix} x^* = x + p \\ \lambda^* \end{pmatrix} \text{ unique solution of KKT system}$$
$$\overset{\text{Theorem 16.2}}{\Rightarrow} x^* \text{ is the unique solution to EQP}$$

- How to solve KKT system (KKT matrix indefinite, but symmetric):
  - Full-space: Symmetric indefinite (LDL) factorization: $P^\top K P = LBL^\top$
  - Reduced space: Use *Ax=b* to eliminate *m* variables. Requires computation of *Z*, basis for nullspace of *A*, which can be costly. Reduced space method can be faster than full-space if *n-m* $\ll$ *n*.

Local SQP-algorithm for solving equality-constrained NLPs

$$\min f(x)$$
$$\text{subject to } c(x) = 0$$

Algorithm 18.1 (Local SQP Algorithm for solving (18.1)).

Choose an initial pair (x_0, λ_0); set $k \leftarrow 0$;

repeat until a convergence test is satisfied

Evaluate $f_k, \nabla f_k, \nabla^2_{xx}\mathcal{L}_k, c_k$, and A_k;

Solve (18.7) to obtain p_k and l_k;

Set $x_{k+1} \leftarrow x_k + p_k$ and $\lambda_{k+1} \leftarrow l_k$;

end (repeat)

EQP:

$$\min_p \quad f_k + \nabla f_k^T p + \tfrac{1}{2} p^T \nabla^2_{xx}\mathcal{L}_k p$$
$$\text{subject to} \quad A_k p + c_k = 0.$$