Numerical Nonlinear Optimization Part III

Andreas Wächter

Center for Nonlinear Studies

July 6, 2020

Outline

Last week:

- Line search and trust region methods for unconstrained optimization.
- Started discussion of optimality conditions for constrained optimization.

Today:

- Optimality conditions for constrained optimization.
- Solving quadratic problems with equality constraints
- Solving quadratic problems with inequality constraints

Next week:

- Sequential Quadratic Programming
- Interior-Point Methods

Constrained Nonlinear Optimization Problems

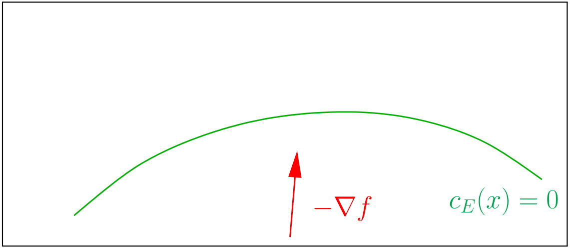
$$\min_{x\in\mathbb{R}^n} f(x)$$
$$\text{s.t. } c_E(x) = 0$$
$$c_I(x) \leq 0$$

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$
$$c_E : \mathbb{R}^n \longrightarrow \mathbb{R}^{n_E}$$
$$c_I : \mathbb{R}^n \longrightarrow \mathbb{R}^{n_I}$$

- We assume that all functions are twice continuously differentiable.
- Often called “Nonlinear Program” (NLP).
- For problems with convex objective and linear equality and convex inequality constraints, every local minimizer is a global minimizer.

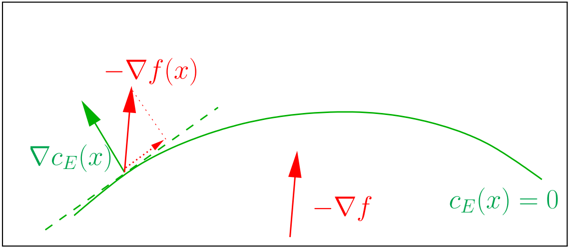
Optimality Conditions: Equality Constraints

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$



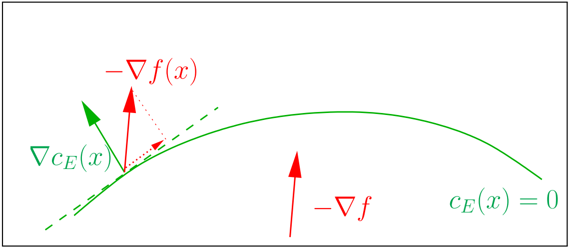
Optimality Conditions: Equality Constraints

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$



Optimality Conditions: Equality Constraints

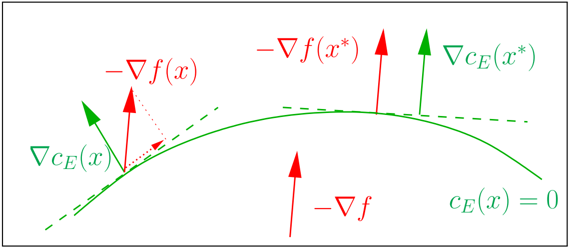
$$\min_{x \in \mathbb{R}^n} \; f(x)$$
$$s.t. \; c_E(x) = 0$$



- Moving along projection of $-\nabla f(x)$ onto tangent space of feasible set decreases objective.

Optimality Conditions: Equality Constraints

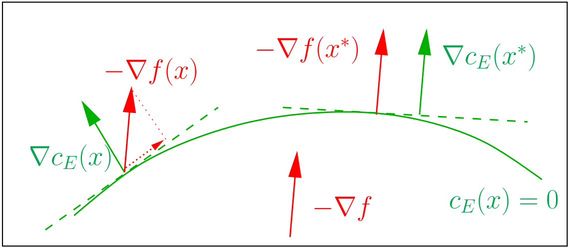
$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$



- Moving along projection of $-\nabla f(x)$ onto tangent space of feasible set decreases objective.

Optimality Conditions: Equality Constraints

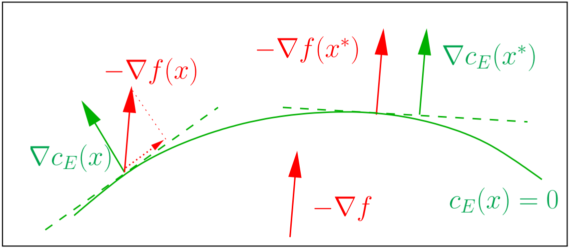
$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$



- Moving along projection of $-\nabla f(x)$ onto tangent space of feasible set decreases objective.
- At local minimum, projection of $-\nabla f(x)$ must be zero.

Optimality Conditions: Equality Constraints

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$

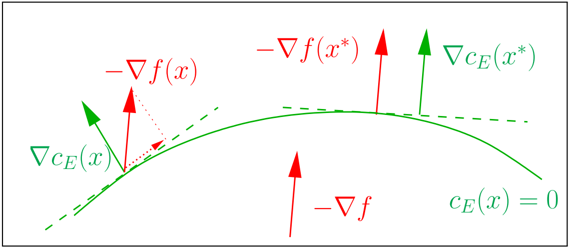


- Moving along projection of $-\nabla f(x)$ onto tangent space of feasible set decreases objective.
- At local minimum, projection of $-\nabla f(x)$ must be zero.
- For this, $-\nabla f(x^*)$ must be linear combination of constraint gradient:

$$\boxed{-\nabla f(x^*) = \nabla c_E(x^*)\, \lambda_E} \qquad \lambda_E \in \mathbb{R}$$

Optimality Conditions: Equality Constraints

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$

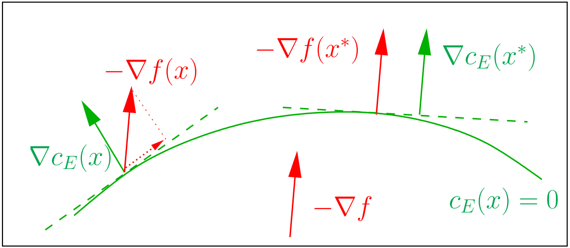


- Moving along projection of $-\nabla f(x)$ onto tangent space of feasible set decreases objective.
- At local minimum, projection of $-\nabla f(x)$ must be zero.
- For this, $-\nabla f(x^*)$ must be linear combination of constraint gradients:

$$-\nabla f(x^*) = \sum_{j=1}^{n_E} \nabla c_{E,j}(x^*)\,\lambda_{E,j} \qquad \lambda_E \in \mathbb{R}^{n_E}$$

Optimality Conditions: Equality Constraints

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$



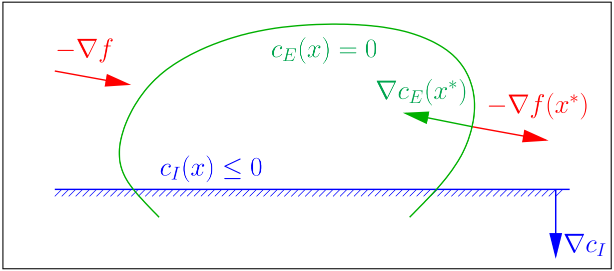
- Moving along projection of $-\nabla f(x)$ onto tangent space of feasible set decreases objective.
- At local minimum, projection of $-\nabla f(x)$ must be zero.
- For this, $-\nabla f(x^*)$ must be linear combination of constraint gradients:

$$-\nabla f(x^*) = \sum_{j=1}^{n_E} \nabla c_{E,j}(x^*)\, \lambda_{E,j} = \nabla c_E(x^*)\, \lambda_E \qquad \lambda_E \in \mathbb{R}^{n_E}$$

- Notation: Columns of $\nabla c_E(x^*)$ are the constraints gradients.

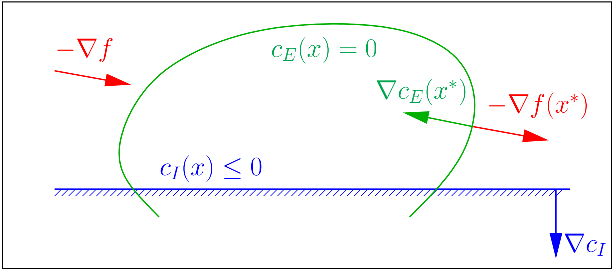
Optimality Conditions: Inequality Constraints

$$\min_{x\in\mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$
$$c_I(x) \le 0$$



Optimality Conditions: Inequality Constraints

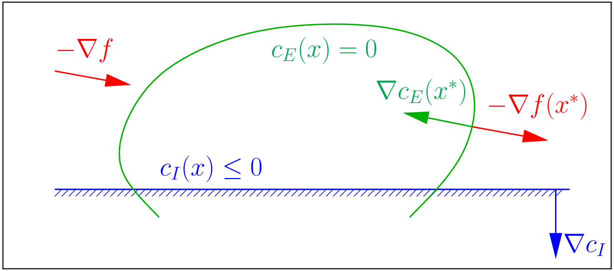
$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$
$$c_I(x) \leq 0$$



- First local minimum:
  - Inequality constraint is inactive (not binding), it might as well not be there.

Optimality Conditions: Inequality Constraints

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$
$$c_I(x) \leq 0$$

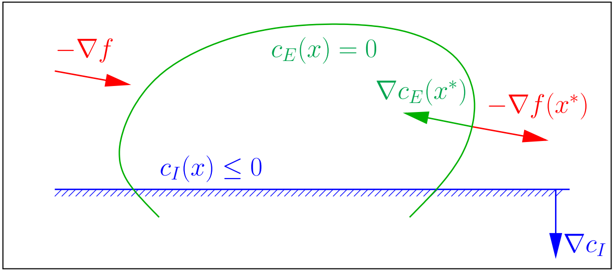


- First local minimum:
  - Inequality constraint is inactive (not binding), it might as well not be there.
- Same relationship as before:

$$-\nabla f(x^*) = \nabla c_E(x^*) \cdot \lambda_E \qquad \lambda_E \in \mathbb{R}$$

Optimality Conditions: Inequality Constraints

$$\min_{x\in\mathbb{R}^n} \; f(x) \quad s.t. \; c_E(x) = 0, \; c_I(x) \le 0$$

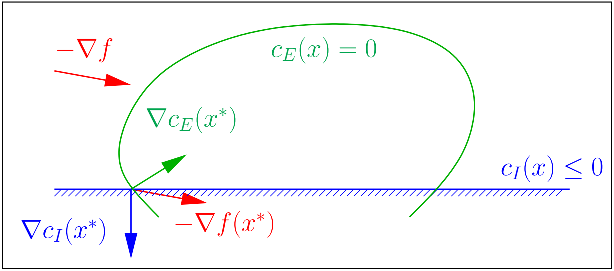


- First local minimum:
  - Inequality constraint is inactive (not binding), it might as well not be there.
- Same relationship as before:

$$-\nabla f(x^*) = \nabla c_E(x^*) \cdot \lambda_E + \nabla c_I(x^*) \cdot \lambda_I \qquad \lambda_E \in \mathbb{R}, \; \lambda_I = 0$$

Optimality Conditions: Inequality Constraints

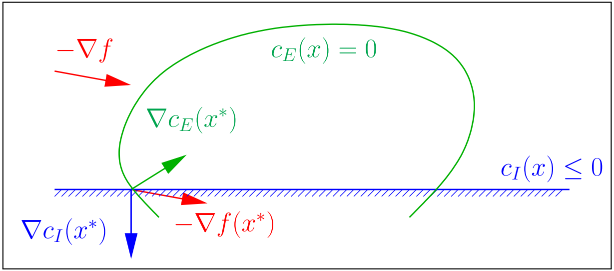
$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$
$$c_I(x) \leq 0$$



- Second local minimum:
  - Inequality constraint is active.

Optimality Conditions: Inequality Constraints

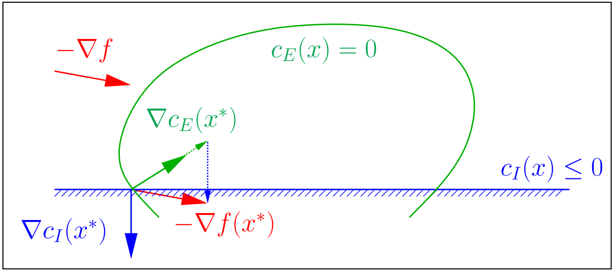
$$\min_{x \in \mathbb{R}^n} \; f(x)$$
$$s.t. \quad c_E(x) = 0$$
$$c_I(x) \leq 0$$



- Second local minimum:
  - Inequality constraint is active.
- Projection of $-\nabla f(x^*)$ onto tangent space of "$c_E(x) = 0$" points into direction that violates "$c_I(x) \leq 0$".

Optimality Conditions: Inequality Constraints

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$
$$c_I(x) \le 0$$

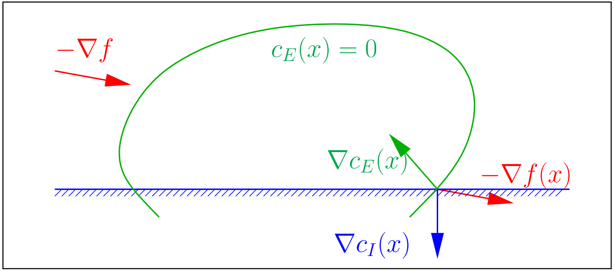


- Second local minimum:
  - Inequality constraint is active.
- Projection of $-\nabla f(x^*)$ onto tangent space of "$c_E(x) = 0$" points into direction that violates "$c_I(x) \le 0$".

$$-\nabla f(x^*) = \nabla c_E(x^*) \cdot \lambda_E + \nabla c_I(x^*) \cdot \lambda_I \qquad \lambda_E \in \mathbb{R}, \ \lambda_I \ge 0$$

Optimality Conditions: Inequality Constraints

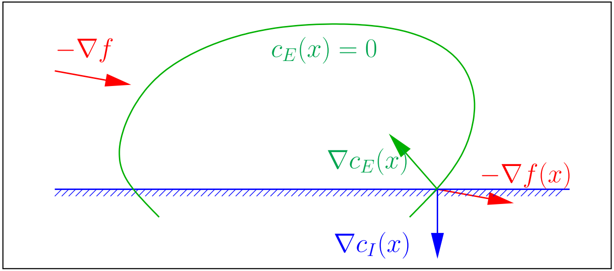
$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$
$$c_I(x) \leq 0$$



- Another point where inequality is active.

Optimality Conditions: Inequality Constraints

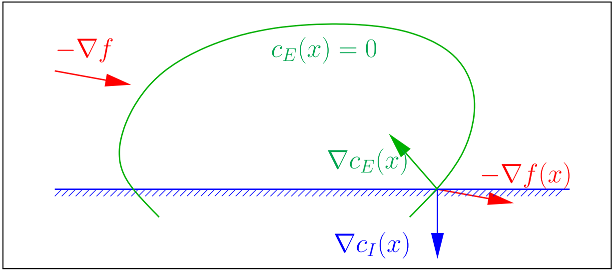
$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$
$$c_I(x) \le 0$$



- Another point where inequality is active.
- Projection of $-\nabla f(x)$ onto tangent space of "$c_E(x) = 0$" points into direction that satisfies "$c_I(x) \le 0$".

Optimality Conditions: Inequality Constraints

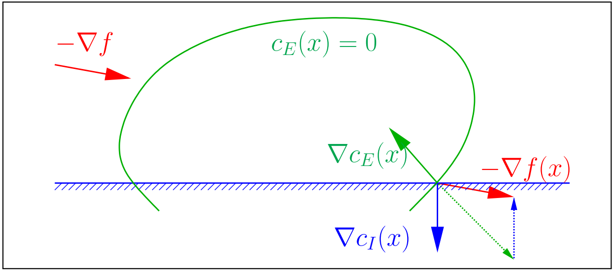
$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$
$$c_I(x) \leq 0$$



- Another point where inequality is active.
- Projection of $-\nabla f(x)$ onto tangent space of "$c_E(x) = 0$" points into direction that satisfies "$c_I(x) \leq 0$".
  - Can move into this direction and improve objective.

Optimality Conditions: Inequality Constraints

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$s.t. \quad c_E(x) = 0$$
$$c_I(x) \leq 0$$



- Another point where inequality is active.
- Projection of $-\nabla f(x)$ onto tangent space of "$c_E(x) = 0$" points into direction that satisfies "$c_I(x) \leq 0$".
  - Can move into this direction and improve objective.

$$-\nabla f(x) = \nabla c_E(x) \cdot \lambda_E + \nabla c_I(x) \cdot \lambda_I \qquad \lambda_E \in \mathbb{R}, \ \lambda_I < 0$$

Summary of Conditions

- Projection of $-\nabla f(x^*)$ onto the right tangent space must be zero:

$$\boxed{\nabla f(x^*) + \nabla c_E(x^*)\lambda_E + \nabla c_I(x^*)\lambda_I = 0}$$

  for some Lagrangian multipliers $\lambda_E \in \mathbb{R}^{n_E}$ and $\lambda_I \in \mathbb{R}^{n_I}$.
  - There is no direction that decreases objective and stays feasible.
- Releasing active inequality does not make it possible to improve objective:

$$\boxed{\lambda_I \geq 0}$$

- Only active constraints can contribute to the (local) optimality conditions:

$$\boxed{c_{I,j}(x^*) \cdot \lambda^*_{I,j} = 0} \quad \text{for all } j = 1, \ldots, n_I$$

  - If constraint is not active, multiplier must be zero.
  - This is called complementarity condition.
  - "At least one of $c_{I,j}(x^*)$ and $\lambda^*_{I,j}$ has to be zero."

KKT Conditions

Theorem (First-Order Necessary Optimality Conditions)

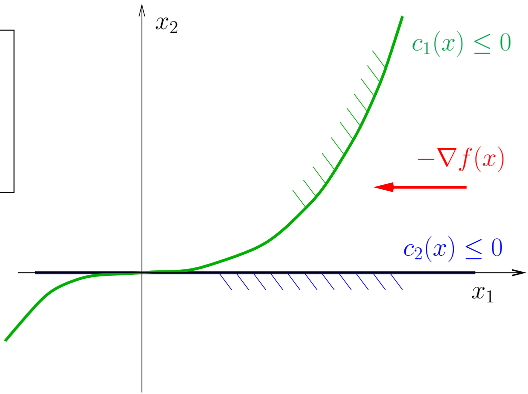
Let x^* be a local minimizer and suppose that f, c_E, and c_I are continuously differentiable. Further assume that a "constraint qualification" holds. Then there exist Lagrangian multipliers $\lambda_E^* \in \mathbb{R}^{n_E}$ and $\lambda_I^* \in \mathbb{R}^{n_I}$ so that the following conditions hold:

$$
\begin{aligned}
\nabla f(x^*) + \nabla c_E(x^*)\lambda_E^* + \nabla c_I(x^*)\lambda_I^* &= 0 \\
c_E(x^*) &= 0 \\
c_I(x^*) &\le 0 \\
\lambda_I^* &\ge 0 \\
c_{I,j}(x^*) \cdot \lambda_{I,j}^* &= 0 \quad \textit{for all } j = 1, \ldots, n_I
\end{aligned}
$$

- These conditions are called the KKT conditions.
  - Named after Karush, Kuhn, and Tucker.

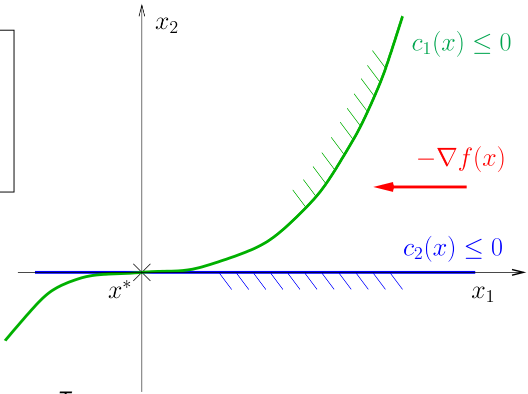
Existence of Multipliers

$$\min_{x \in \mathbb{R}^2} \quad f(x) = x_1$$
$$s.t. \quad c_1(x) = x_2 - x_1^3 \leq 0$$
$$c_2(x) = -x_2 \leq 0$$



Existence of Multipliers

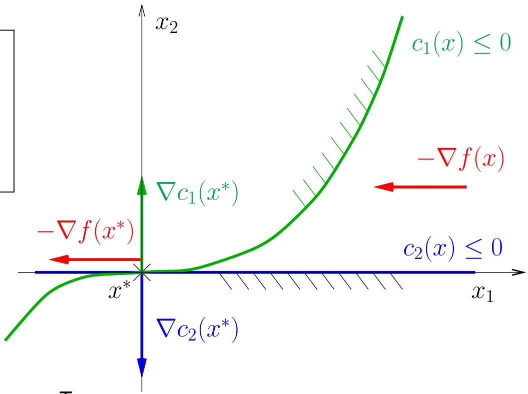
$$\min_{x \in \mathbb{R}^2} \quad f(x) = x_1$$
$$s.t. \quad c_1(x) = x_2 - x_1^3 \leq 0$$
$$c_2(x) = -x_2 \leq 0$$



- Optimal solution: $x^* = (0, 0)^T$

Existence of Multipliers

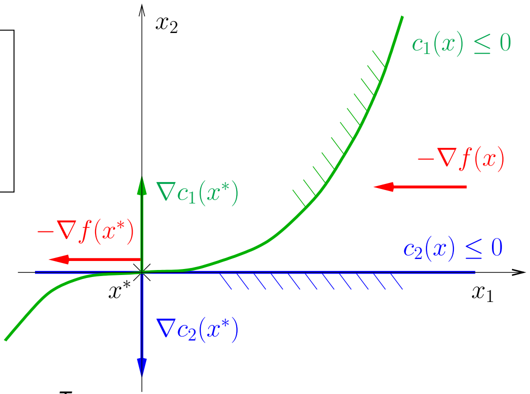
$$\min_{x \in \mathbb{R}^2} \quad f(x) = x_1$$
$$s.t. \quad c_1(x) = x_2 - x_1^3 \leq 0$$
$$c_2(x) = -x_2 \leq 0$$



- Optimal solution: $x^* = (0, 0)^T$

Existence of Multipliers

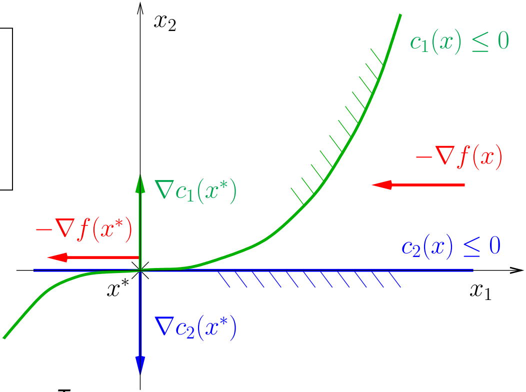
$$\min_{x \in \mathbb{R}^2} \quad f(x) = x_1$$
$$s.t. \quad c_1(x) = x_2 - x_1^3 \le 0$$
$$c_2(x) = -x_2 \le 0$$



- Optimal solution: $x^* = (0,0)^T$
- $-\nabla f(x^*)$ is not a linear combination of constraint gradients!

Existence of Multipliers

$$\min_{x \in \mathbb{R}^2} \quad f(x) = x_1$$
$$s.t. \quad c_1(x) = x_2 - x_1^3 \leq 0$$
$$c_2(x) = -x_2 \leq 0$$



- Optimal solution: $x^* = (0, 0)^T$
- $-\nabla f(x^*)$ is not a linear combination of constraint gradients!
- No Lagrangian multipliers exist.

Constraint Qualifications

- A constraint qualification is a condition that ensures the existence of Lagrangian multipliers.
- If no multipliers exist, algorithms that seek KKT points might have difficulties or fail!
- Ipopt heuristic: "$c_I(x) \leq$ bound_relax_factor"
  - Relaxed solution more likely to satisfy constraint qualification.

Examples:

- Linear-Independence Constraint Qualification (LICQ)
  - The constraint gradients for all active constraints are linearly independent.
- All constraints are linear, e.g., Linear Programs.
- Mangasarian-Fromovitz Constraint Qualification (MFCQ)
  - Looser than LICQ.

Lagrangian Function

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x) \\ s.t. \quad & c_E(x) = 0 \\ & c_I(x) \leq 0 \end{aligned} \qquad \text{(NLP)}$$

- The Lagrangian function for (NLP) is defined as

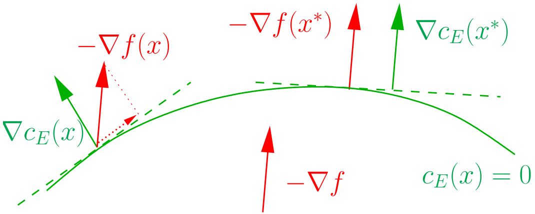
$$\mathcal{L}(x, \lambda_E, \lambda_I) = f(x) + c_E(x)^T \lambda_E + c_I(x)^T \lambda_I$$

- Helps to express relationships and optimality conditions.
- For example, first equation in KKT conditions:

$$0 = \nabla f(x^*) + \nabla c_E(x^*)\lambda_E^* + \nabla c_I(x^*)\lambda_I^* = \nabla_x \mathcal{L}(x^*, \lambda_E^*, \lambda_I^*)$$

Null Space of Constraint Gradients

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$\text{s.t.} \quad c_E(x) = 0$$

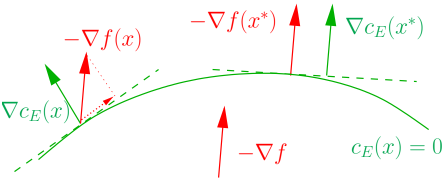


- It only matters how the objective changes within the feasible set.
- Look at directions in the null space of constraint gradients:

$$N_\Omega(x^*) = \{d \in \mathbb{R}^n : \nabla c_E(x^*)^T d = 0\}$$

Second-Order Optimality Conditions For Equality-Constrained Problems

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$\text{s.t.} \quad c_E(x) = 0$$



- Hessian of Lagrangian function

$$\nabla^2_{xx}\mathcal{L}(x^*, \lambda^*_E) = \nabla^2 f(x^*) + \sum_{j=1}^{n_E} \nabla^2 c_{E,j}(x^*) \cdot \lambda^*_{E,j}$$

captures curvature of objective and constraints.

- Necessary second-order optimality condition:

$$d^T \nabla^2_{xx}\mathcal{L}(x^*, \lambda^*_E) d \geq 0 \text{ for all } d \in N_\Omega(x^*)$$

Strict Complementarity

Definition (Strict Complementarity)

Let x^* a local minimizer and λ_E^* and λ_I^* be Lagrangian multipliers so that the KKT conditions hold. We say that strict complementarity holds if

$$\boxed{c_{I,j}(x^*) < 0 \quad \text{or} \quad \lambda_{I,j} > 0 \quad \text{for all } j = 1, \ldots, n_I}$$

- If an inequality is active, its multiplier is non-zero.
- Then the inequality constraint is “strongly binding”; we can treat it as equality constraint in the 2nd-order optimality conditions.

Null Space of Active Constraints

Active set:

- A constraint that holds with equality at $x \in \Omega$ is "active at x".
- Active set $\mathcal{A}(x)$ for $x \in \Omega$:
  - Indices of all constraints that are active at x, including all c_E.

Null space of active constraint gradients:

$$N_\Omega(x^*) = \{d \in \mathbb{R}^n : \nabla c_j(x^*)^T d = 0 \text{ for all } j \in \mathcal{A}(x^*)\}$$

Necessary Second-Order Optimality Conditions

Theorem (Necessary Second-Order Optimality Conditions)

Let x^* be a local minimizer with KKT multipliers λ_E^* and λ_I^* at which LICQ and strict complementarity holds. Then

$$d^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda_E^*, \lambda_I^*) d \geq 0 \text{ for all } d \in N_\Omega(x^*)$$

Theorem (Sufficient Second-Order Optimality Conditions)

Let x^*, λ_E^*, and λ_I^* be such that the KKT conditions and strict complementarity holds. If

$$d^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda_E^*, \lambda_I^*) d > 0 \text{ for all } d \in N_\Omega(x^*) \setminus \{0\}$$

then x^* is a strict local minimizer.

Quadratic Programming

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & \tfrac{1}{2}x^TQx + g^Tx \\ \text{s.t.} \quad & A_E x + b_E = 0 \\ & A_I x + b_I \le 0 \end{aligned} \qquad \text{(QP)}$$

$$Q \in \mathbb{R}^{n\times n} \text{ symmetric}$$
$$A_E \in \mathbb{R}^{n_E\times n} \qquad b_E \in \mathbb{R}^{n_E}$$
$$A_I \in \mathbb{R}^{n_I\times n} \qquad b_I \in \mathbb{R}^{n_I}$$

- Many applications (e.g., portfolio optimization, optimal control).
- Important building block for methods for general NLP.
- Algorithms:
  - Active-set methods
  - Interior-point methods
- Let's first consider equality-constrained case.
- Assume: all rows of A_E are linearly independent.

Equality-Constrained QP

$$\min_{x\in\mathbb{R}^n} \; \frac{1}{2}x^T Q x + g^T x \qquad \text{(EQP)}$$
$$\text{s.t. } Ax + b = 0$$

First-order optimality conditions:

$$Qx + g + A^T\lambda = 0$$
$$Ax + b = 0$$

Find stationary point (x^*, λ^*) by solving the linear system

$$\begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix} \begin{pmatrix} x^* \\ \lambda^* \end{pmatrix} = -\begin{pmatrix} g \\ b \end{pmatrix}.$$

KKT System of QP

$$\begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix} \begin{pmatrix} x^* \\ \lambda^* \end{pmatrix} = - \begin{pmatrix} g \\ b \end{pmatrix}$$

- When is (x^*, λ^*) indeed a solution of (EQP)?
- Recall the sufficient second-order optimality condition:
  - If KKT conditions and

$$d^T Q d > 0 \text{ for all } d \in N_\Omega(x^*) \setminus \{0\}$$

hold, then x^* is a strict local minimizer of (EQP).

- On the other hand:
  - If Q has negative eigenvalue in $N_\Omega(x^*)$, then (EQP) is unbounded below.

Direct Solution of the KKT System

$$\underbrace{\begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix}}_{=:K} \begin{pmatrix} x^* \\ \lambda^* \end{pmatrix} = -\begin{pmatrix} g \\ b \end{pmatrix}$$

- Can we verify that x^* is minimizer without computing $N_\Omega(x^*)$?

Definition (Inertia of Matrix)

Let n_+, n_-, n_0 be the number of positive, negative, and zero eigenvalues of a symmetric matrix K. Then $\text{In}(K) = (n_+, n_-, n_0)$ is called the inertia of K.

Theorem

Suppose that A has full rank. If $\text{In}(K) = (n, n_E, 0)$, then x^* is the unique global minimizer of (EQP).

Computing the Inertia

$$\underbrace{\begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix}}_{=:K} \begin{pmatrix} x^* \\ \lambda^* \end{pmatrix} = - \begin{pmatrix} g \\ b \end{pmatrix}$$

- Symmetric indefinite factorization $K = LBL^T$
  - L: unit lower triangular matrix
  - B: block diagonal matrix with 1×1 and 2×2 diagonal blocks
- Can be computed efficiently, exploits sparsity.
- Factorization used to solve the linear system.
- Obtain inertia from counting eigenvalues of the blocks in B.
  - This is easy!

Ways to Solve Equality-Constrained QPs

- Direct method:
  - Factorize KKT matrix.
  - If $L^T BL$ factorization is used, we can determine if x^* is indeed a minimizer.
  - Easy general-purpose option.
- Schur-complement method:
  - Requires that Q is positive definite and easy to factorize (e.g., diagonal).
  - Number of constraints n_E should not be large.
  - Often used in interior-point LP solvers.
- Null-space method:
  - Step decomposition into range-space step and null-space step.
  - Permits exploitation of constraint matrix structure.
  - Number of degrees of freedom $(n - n_E)$ should not be large.

Inequality-Constrained QPs

$$\min_{x\in\mathbb{R}^n} \tfrac{1}{2}x^TQx + g^Tx$$
$$\text{s.t. } a_i^Tx + b_i = 0 \text{ for } i \in \mathcal{E}$$
$$a_i^Tx + b_i \le 0 \text{ for } i \in \mathcal{I}$$

$$Qx + g + \sum_{i\in\mathcal{E}\cup\mathcal{I}} a_i\lambda_i = 0$$
$$a_i^Tx + b_i = 0 \text{ for } i \in \mathcal{E}$$
$$a_i^Tx + b_i \le 0 \text{ for } i \in \mathcal{I}$$
$$\lambda_i \ge 0 \text{ for } i \in \mathcal{I}$$
$$(a_i^Tx + b_i)\lambda_i = 0 \text{ for } i \in \mathcal{I}$$

- Assume here:
  - Q is positive definite.
  - $\{a_i\}_{i\in\mathcal{E}}$ are linearly independent.
- Difficulty: Decide, which inequality constraints are active at x^*.
- If that was known, could just solve equality-constrained QPs.

Working Set

Choose working set $\mathcal{W} \subseteq \mathcal{I}$ (guess of optimal active set) and solve

(EQP)

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \ & \tfrac{1}{2} x^T Q x + g^T x \\ \text{s.t. } & a_i^T x + b_i = 0 \text{ for } i \in \mathcal{E} \\ & a_i^T x + b_i = 0 \text{ for } i \in \mathcal{W} \end{aligned}$$

$$\begin{aligned} Qx + g + \sum_{i \in \mathcal{E} \cup \mathcal{W}} a_i \lambda_i &= 0 \\ a_i^T x + b_i &= 0 \text{ for } i \in \mathcal{E} \\ a_i^T x + b_i &= 0 \text{ for } i \in \mathcal{W} \end{aligned}$$

Solution of KKT system for (EQP) gives

$$x^{\text{EQP}} \in \mathbb{R}^n \text{ and } \lambda_i^{\text{EQP}} \text{ for } i \in \mathcal{E} \cup \mathcal{W}$$

Complete to candidate optimal KKT solution we set

$$\lambda_i^{\text{EQP}} = 0 \text{ for } i \in \mathcal{I} \setminus \mathcal{W}$$

Optimality Test

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}x^TQx + g^Tx$$
$$\text{s.t. } a_i^Tx + b_i = 0 \text{ for } i \in \mathcal{E}$$
$$a_i^Tx + b_i = 0 \text{ for } i \in \mathcal{W}$$

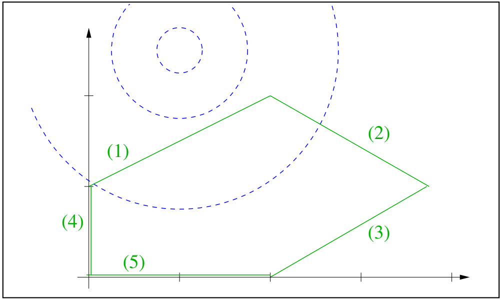
$$Qx + g + \sum_{i\in\mathcal{E}\cup\mathcal{W}} a_i\lambda_i = 0$$
$$a_i^Tx + b_i = 0 \text{ for } i \in \mathcal{E}$$
$$a_i^Tx + b_i = 0 \text{ for } i \in \mathcal{W}$$

Check if $(x^{\text{EQP}}, \lambda^{\text{EQP}})$ is optimal KKT point for (QP):

$$a_i^Tx^{\text{EQP}} + b_i \overset{?}{\leq} 0 \text{ for } i \in \mathcal{I}\setminus\mathcal{W}$$
$$\lambda_i^{\text{EQP}} \overset{?}{\geq} 0 \text{ for } i \in \mathcal{I}$$

- Complementarity holds by construction ($\lambda_i = 0$ for $i \in \mathcal{I}\setminus\mathcal{W}$).
- If satisfied, $(x^{\text{EQP}}, \lambda^{\text{EQP}})$ is the (unique) optimal solution.
- Otherwise, let's try a different working set.

Demonstration on Example QP



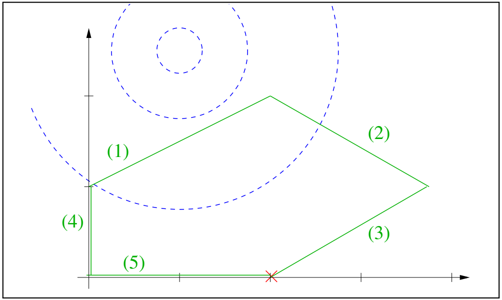
$$\min\ (x_1 - 1)^2 + (x_2 - 2.5)^2$$

$$\text{s.t.}\ -x_1 + 2x_2 - 2 \le 0\ (1) \qquad -x_1 \le 0\ (4)$$

$$x_1 + 2x_2 - 6 \le 0\ (2) \qquad -x_2 \le 0\ (5)$$

$$x_1 - 2x_2 - 2 \le 0\ (3)$$

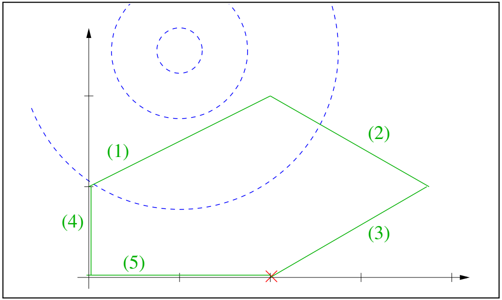
Primal Active-Set QP Solver Iteration 1



Initialization:
Choose feasible starting iterate x

$x = (0, 2)$

Primal Active-Set QP Solver Iteration 1



$$\mathcal{W} = \{3, 5\}$$
$$x = (0, 2)$$

Initialization:

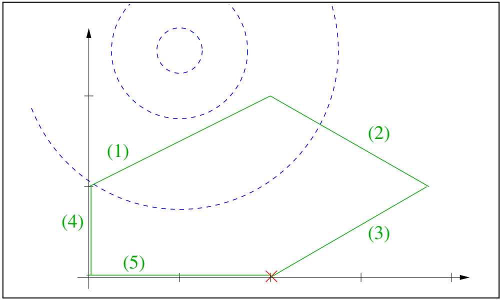
Choose feasible starting iterate x

Choose working set $\mathcal{W} \subseteq \mathcal{I}$ with

- $i \in \mathcal{W} \Longrightarrow a_i^T x + b_i = 0$
- $\{a_i\}_{i \in \mathcal{E} \cup \mathcal{W}}$ are linear independent

(Algorithm will maintain these properties)

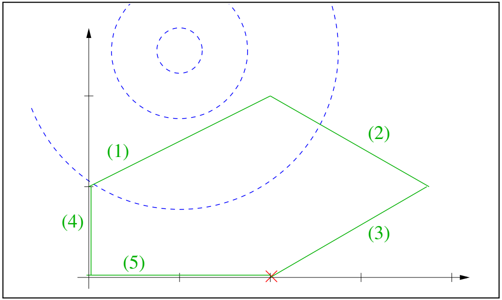
Primal Active-Set QP Solver Iteration 1



$$\mathcal{W} = \{3, 5\}$$
$$x = (0, 2)$$
$$x^{\text{EQP}} = (0, 2)$$
$$\lambda_3 = -2$$
$$\lambda_5 = -1$$

Solve (EQP)

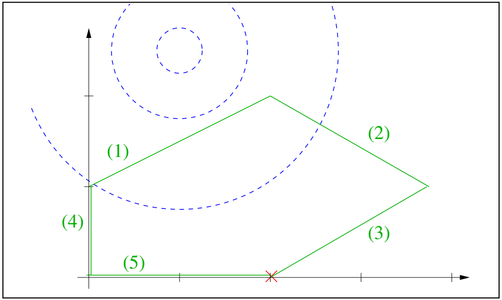
Primal Active-Set QP Solver Iteration 1



$$\mathcal{W} = \{3, 5\}$$
$$x = (0, 2)$$
$$x^{\text{EQP}} = (0, 2)$$
$$\lambda_3 = -2$$
$$\lambda_5 = -1$$

Status: Current iterate is optimal for (EQP).

Primal Active-Set QP Solver Iteration 1



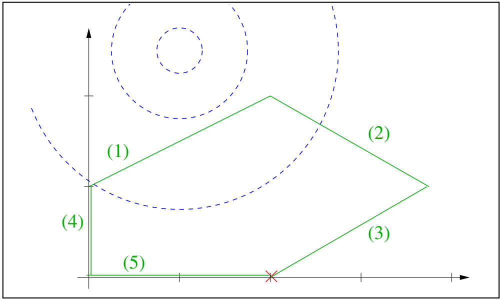
$$\mathcal{W} = \{3, 5\}$$
$$x = (0, 2)$$
$$x^{\text{EQP}} = (0, 2)$$
$$\lambda_3 = -2$$
$$\lambda_5 = -1$$

Status: Current iterate is optimal for (EQP).

Release Constraint:

- Pick constraint i with $\lambda_i < 0$ (here $i = 3$).

Primal Active-Set QP Solver Iteration 1



$$\mathcal{W} = \{3, 5\}$$
$$x = (0, 2)$$
$$x^{\text{EQP}} = (0, 2)$$
$$\lambda_3 = -2$$
$$\lambda_5 = -1$$

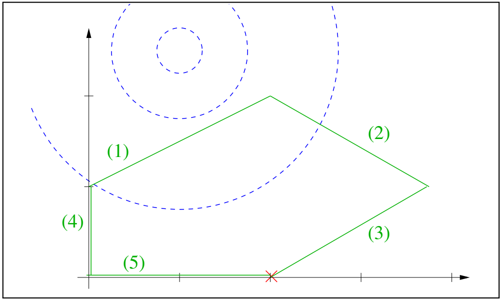
Status: Current iterate is optimal for (EQP).

Release Constraint:

- Pick constraint i with $\lambda_i < 0$ (here $i = 3$).
- Remove i from working set:

$$\mathcal{W} \leftarrow \mathcal{W} \setminus \{3\} = \{5\}$$

Primal Active-Set QP Solver Iteration 1



$$\mathcal{W} = \{3, 5\}$$
$$x = (0, 2)$$
$$x^{\text{EQP}} = (0, 2)$$
$$\lambda_3 = -2$$
$$\lambda_5 = -1$$

Status: Current iterate is optimal for (EQP).

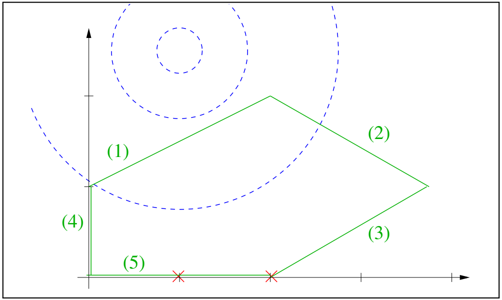
Release Constraint:

- Pick constraint i with $\lambda_i < 0$ (here $i = 3$).
- Remove i from working set:

$$\mathcal{W} \leftarrow \mathcal{W} \setminus \{3\} = \{5\}$$

- Keep iterate $x = (0, 2)$.

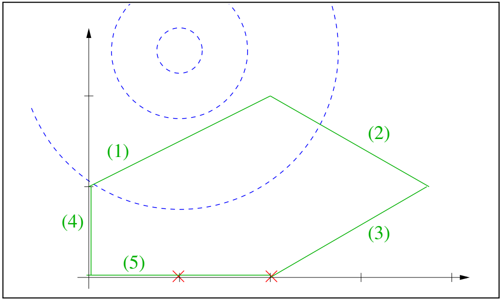
Primal Active-Set QP Solver Iteration 2



$$\mathcal{W} = \{5\}$$
$$x = (2, 0)$$
$$x^{\text{EQP}} = (1, 0)$$
$$\lambda_5 = -5$$

Solve (EQP)

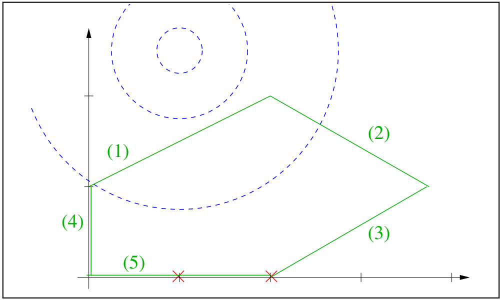
Primal Active-Set QP Solver Iteration 2



$$\mathcal{W} = \{5\}$$
$$x = (2, 0)$$
$$x^{\text{EQP}} = (1, 0)$$
$$\lambda_5 = -5$$

Status: Current iterate is not optimal for (EQP).

Primal Active-Set QP Solver Iteration 2



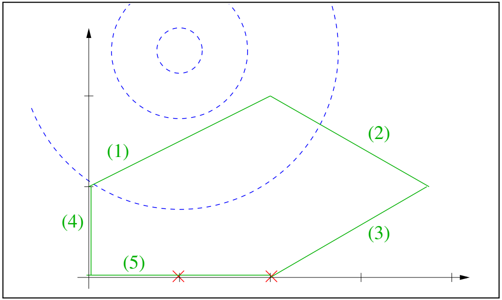
$$\mathcal{W} = \{5\}$$
$$x = (2, 0)$$
$$x^{\text{EQP}} = (1, 0)$$
$$\lambda_5 = -5$$

Status: Current iterate is not optimal for (EQP).

Take step (x^{EQP} is feasible for original QP):

- Update iterate $x \leftarrow x^{\text{EQP}}$

Primal Active-Set QP Solver Iteration 2



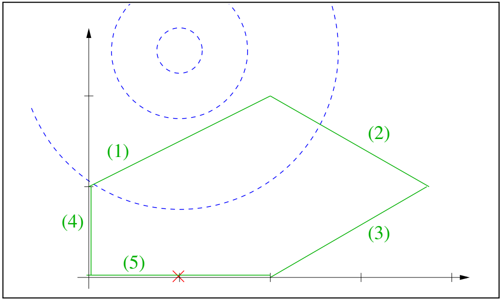
$$\mathcal{W} = \{5\}$$
$$x = (2, 0)$$
$$x^{\text{EQP}} = (1, 0)$$
$$\lambda_5 = -5$$

Status: Current iterate is not optimal for (EQP).

Take step (x^{EQP} is feasible for original QP):

- Update iterate $x \leftarrow x^{\text{EQP}}$
- Keep $\mathcal{W}$

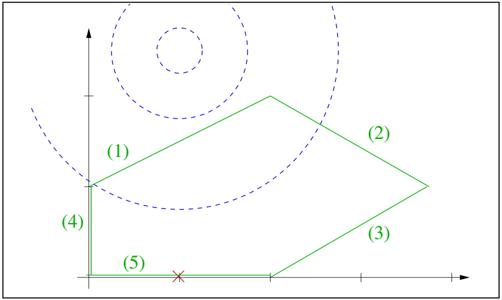
Primal Active-Set QP Solver Iteration 3



$$\mathcal{W} = \{5\}$$
$$x = (1, 0)$$
$$x^{\text{EQP}} = (1, 0)$$
$$\lambda_5 = -5$$

Solve (EQP)

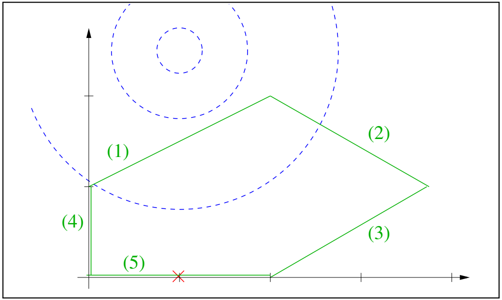
Primal Active-Set QP Solver Iteration 3



Status: Current iterate is optimal for (EQP)

$$\mathcal{W} = \{5\}$$
$$x = (1, 0)$$
$$x^{\text{EQP}} = (1, 0)$$
$$\lambda_5 = -5$$

Primal Active-Set QP Solver Iteration 3



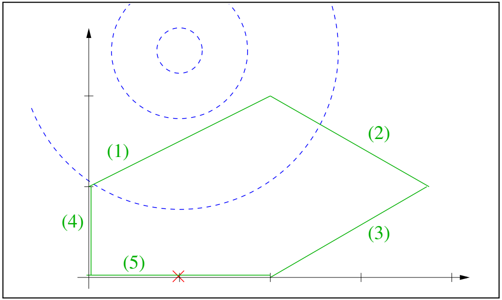
Status: Current iterate is optimal for (EQP)

$$\mathcal{W} = \{5\}$$
$$x = (1,0)$$
$$x^{\text{EQP}} = (1,0)$$
$$\lambda_5 = -5$$

Release Constraint:

- Pick constraint i with $\lambda_i < 0$ (here $i = 5$).

Primal Active-Set QP Solver Iteration 3



$$\mathcal{W} = \{5\}$$
$$x = (1, 0)$$
$$x^{\text{EQP}} = (1, 0)$$
$$\lambda_5 = -5$$

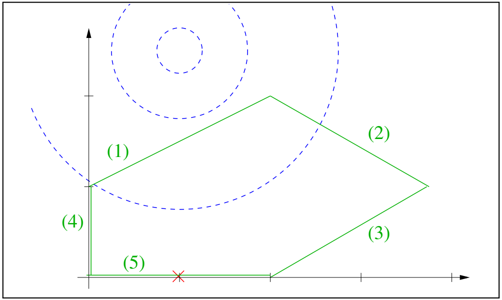
Status: Current iterate is optimal for (EQP)

Release Constraint:

- Pick constraint i with $\lambda_i < 0$ (here $i = 5$).
- Remove i from working set:

$$\mathcal{W} \leftarrow \mathcal{W} \setminus \{5\} = \emptyset$$

Primal Active-Set QP Solver Iteration 3



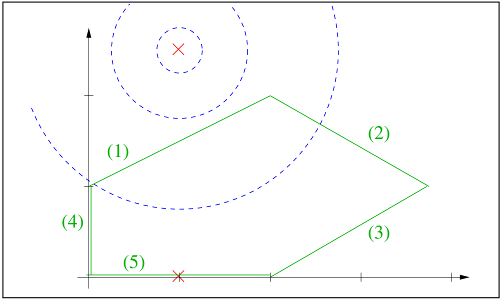
$$\mathcal{W} = \{5\}$$
$$x = (1, 0)$$
$$x^{\text{EQP}} = (1, 0)$$
$$\lambda_5 = -5$$

Status: Current iterate is optimal for (EQP)

Release Constraint:

- Pick constraint i with $\lambda_i < 0$ (here $i = 5$).
- Remove i from working set:
$$\mathcal{W} \leftarrow \mathcal{W} \setminus \{5\} = \emptyset$$
- Keep iterate $x = (1, 0)$.

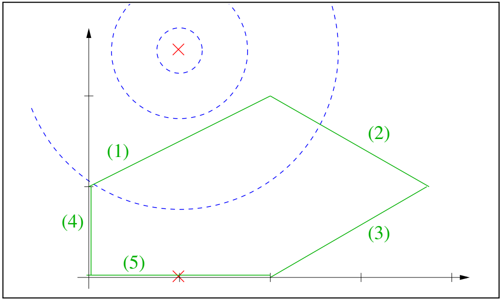
Primal Active-Set QP Solver Iteration 4



$$\mathcal{W} = \emptyset$$
$$x = (1, 0)$$
$$x^{\text{EQP}} = (1, 2.5)$$

Solve (EQP)

Primal Active-Set QP Solver Iteration 4



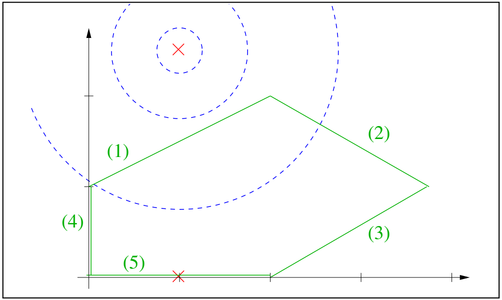
Status: Current iterate not optimal for (EQP)

$$\mathcal{W} = \emptyset$$

$$x = (1, 0)$$

$$x^{\text{EQP}} = (1, 2.5)$$

Primal Active-Set QP Solver Iteration 4

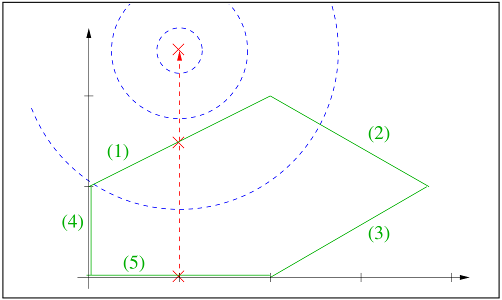


Status: Current iterate not optimal for (EQP)

Take step (x^{EQP} not feasible for original QP):

$$\mathcal{W} = \emptyset$$
$$x = (1, 0)$$
$$x^{\text{EQP}} = (1, 2.5)$$

Primal Active-Set QP Solver Iteration 4



Status: Current iterate not optimal for (EQP)

$\mathcal{W} = \emptyset$

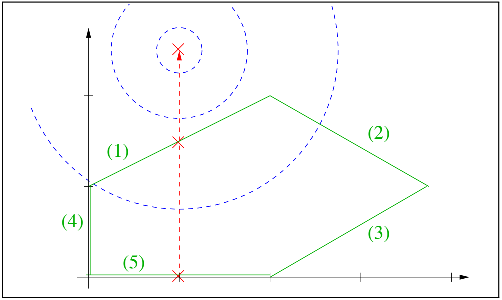
$x = (1, 0)$

$x^{\text{EQP}} = (1, 2.5)$

Take step (x^{EQP} not feasible for original QP):

- Largest $\alpha \in [0, 1]$: $x + \alpha(x^{\text{EQP}} - x)$ feasible

Primal Active-Set QP Solver Iteration 4



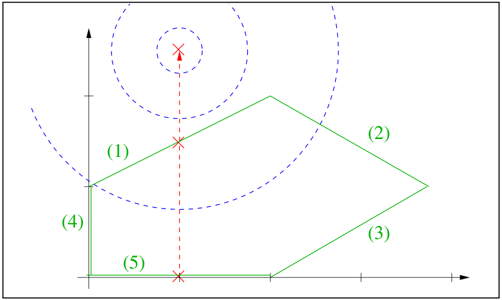
Status: Current iterate not optimal for (EQP)

Take step (x^{EQP} not feasible for original QP):

- Largest $\alpha \in [0,1]$: $x + \alpha(x^{\mathrm{EQP}} - x)$ feasible
- Update iterate $x \leftarrow x + \alpha(x^{\mathrm{EQP}} - x)$

$$\mathcal{W} = \emptyset$$
$$x = (1,0)$$
$$x^{\mathrm{EQP}} = (1,2.5)$$

Primal Active-Set QP Solver Iteration 4



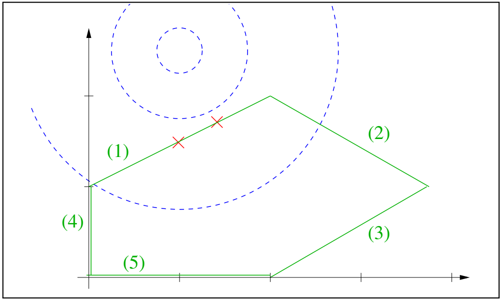
Status: Current iterate not optimal for (EQP)

Take step (x^{EQP} not feasible for original QP):

- Largest $\alpha \in [0,1]$: $x + \alpha(x^{\text{EQP}} - x)$ feasible
- Update iterate $x \leftarrow x + \alpha(x^{\text{EQP}} - x)$
- Update $\mathcal{W} \leftarrow \mathcal{W} \cup \{i\} = \{1\}$
  - where constraint $i = 1$ is "blocking"

$$\mathcal{W} = \emptyset$$
$$x = (1,0)$$
$$x^{\text{EQP}} = (1, 2.5)$$

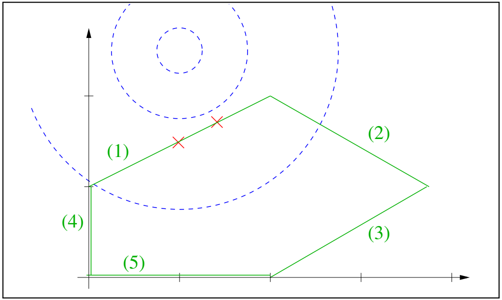
Primal Active-Set QP Solver Iteration 5



$$\mathcal{W} = \{1\}$$
$$x = (1, 1.5)$$
$$x^{\text{EQP}} = (1.4, 1.7)$$
$$\lambda_1 = 0.8$$

Solve (EQP)

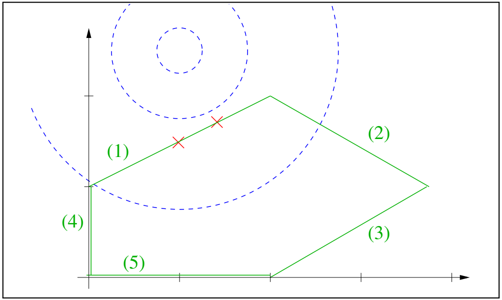
Primal Active-Set QP Solver Iteration 5



$$\mathcal{W} = \{1\}$$
$$x = (1, 1.5)$$
$$x^{\text{EQP}} = (1.4, 1.7)$$
$$\lambda_1 = 0.8$$

Status: Current iterate is not optimal for (EQP).

Primal Active-Set QP Solver Iteration 5



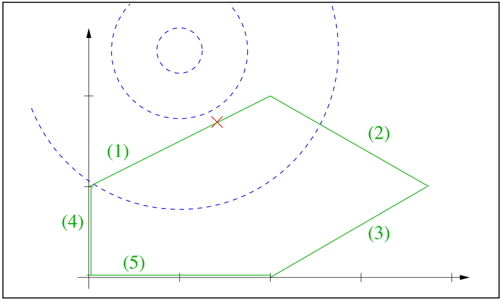
$$\mathcal{W} = \{1\}$$
$$x = (1, 1.5)$$
$$x^{\text{EQP}} = (1.4, 1.7)$$
$$\lambda_1 = 0.8$$

Status: Current iterate is not optimal for (EQP).

Take step (x^{EQP} feasible for original QP):

- Update iterate $x \leftarrow x^{\text{EQP}}$.
- Keep $\mathcal{W}$.

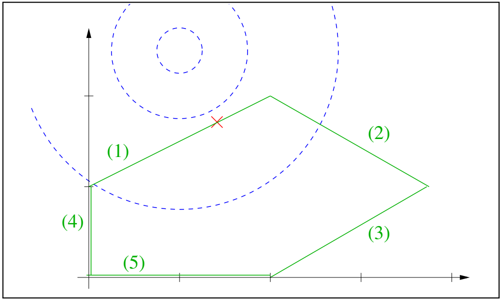
Primal Active-Set QP Solver Iteration 6



$$\mathcal{W} = \{1\}$$
$$x = (1.4, 1.7)$$
$$x^{\text{EQP}} = (1.4, 1.7)$$
$$\lambda_1 = 0.8$$

Solve (EQP)

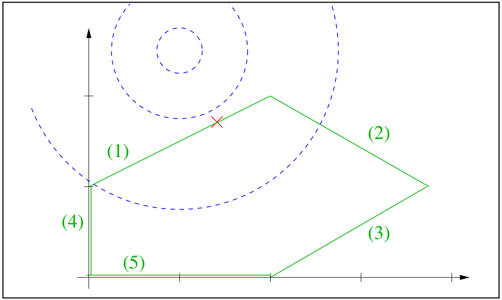
Primal Active-Set QP Solver Iteration 6



$$\mathcal{W} = \{1\}$$
$$x = (1.4, 1.7)$$
$$x^{\text{EQP}} = (1.4, 1.7)$$
$$\lambda_1 = 0.8$$

Status: Current iterate is optimal for (EQP)

Primal Active-Set QP Solver Iteration 6

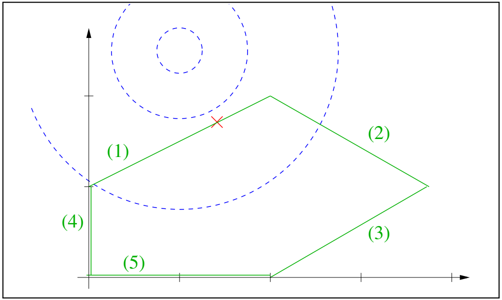


$$\mathcal{W} = \{1\}$$
$$x = (1.4, 1.7)$$
$$x^{\text{EQP}} = (1.4, 1.7)$$
$$\lambda_1 = 0.8$$

Status: Current iterate is optimal for (EQP)

- $\lambda_i \geq 0$ for all $i \in \mathcal{W}$.

Primal Active-Set QP Solver Iteration 6



$$\mathcal{W} = \{1\}$$
$$x = (1.4, 1.7)$$
$$x^{\text{EQP}} = (1.4, 1.7)$$
$$\lambda_1 = 0.8$$

Status: Current iterate is optimal for (EQP)

- $\lambda_i \geq 0$ for all $i \in \mathcal{W}$.

Declare Optimality!

Primal Active-Set QP Method

1: Select feasible x and $\mathcal{W} \subseteq \mathcal{I} \cap \mathcal{A}(x)$.
2: Solve (EQP) to get x^{EQP} and λ^{EQP}.
3: **if** $x = x^{\text{EQP}}$ **then**
4:     If $\lambda^{\text{EQP}} \geq 0$: STOP: Done!
5:     Otherwise, select $\lambda_i^{\text{EQP}} < 0$ and set $\mathcal{W} \leftarrow \mathcal{W} \setminus \{i\}$.
6: **else**
7:     Compute step $p = x^{\text{EQP}} - x$.
8:     Compute $\alpha = \arg\max\{\alpha \in [0,1] : x + \alpha p \text{ is feasible}\}$.
9:     **if** $\alpha < 1$ **then**
10:         Pick $i \in \mathcal{I} \setminus \mathcal{W}$ with $a_i^T p > 0$ and $a_i^T(x + \alpha p) + b_i = 0$.
11:         Set $\mathcal{W} \leftarrow \mathcal{W} \cup \{i\}$.
12:     **end if**
13:     Update $x \leftarrow x + \alpha p$.
14: **end if**
15: Go to step 2.

Primal Active-Set QP Algorithms

- Keeps all iterates feasible.
- Changes $\mathcal{W}$ by at most one constraint per iteration.
- $\{a_i\}_{i \in \mathcal{E} \cup \mathcal{W}}$ remain linearly independent.
- Finite convergence:
  - Finitely many options for $\mathcal{W}$.
  - Objective decreases with every step; as long as $\alpha > 0$!
  - Special handling of degeneracy ($\alpha = 0$ steps) required
- Efficient solution of (EQP)
  - Update the factorization of KKT matrix when $\mathcal{W}$ changes.
- There are variants that allow Q to be indefinite.
- There are other types of active-set methods for QPs.
  - Dual, homotopy, simplex-like, . . .