# MVE165/MMG630, Applied Optimization
# Lecture 13
# Constrained non-linear programming models and algorithms

Ann-Brith Strömberg

2010–04–27

# Constrained nonlinear programming models (Ch. 12)

- The **general model** can be expressed as

$$\begin{aligned} \text{minimize}_{\mathbf{x}\in\Re^n} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & g_i(\mathbf{x}) \leq 0, \quad i \in \mathcal{L}, \\ & h_i(\mathbf{x}) = 0, \quad i \in \mathcal{E}. \end{aligned}$$

- **Convex program:**
  $f$ convex, $g_i$ convex, $i \in \mathcal{L}$, $h_i(\mathbf{x}) = \mathbf{a}_i^{\mathrm{T}}\mathbf{x} - b_i, i \in \mathcal{E}$
  - Any local optimum is a global optimum
- **Quadratic program:**
  $f(\mathbf{x}) = \mathbf{c}^{\mathrm{T}}\mathbf{x} + \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x}$, $g_i(\mathbf{x}) = \mathbf{a}_i^{\mathrm{T}}\mathbf{x} - b_i, i \in \mathcal{L}$,
  $h_i(\mathbf{x}) = \mathbf{k}_i^{\mathrm{T}}\mathbf{x} - \ell_i, i \in \mathcal{E}$
  - The KKT conditions lead to a linear system of inequalities + complementarity

# An algorithm for minimizing a convex function over a bounded polyhedron (Frank–Wolfe) (Ch.12.1)

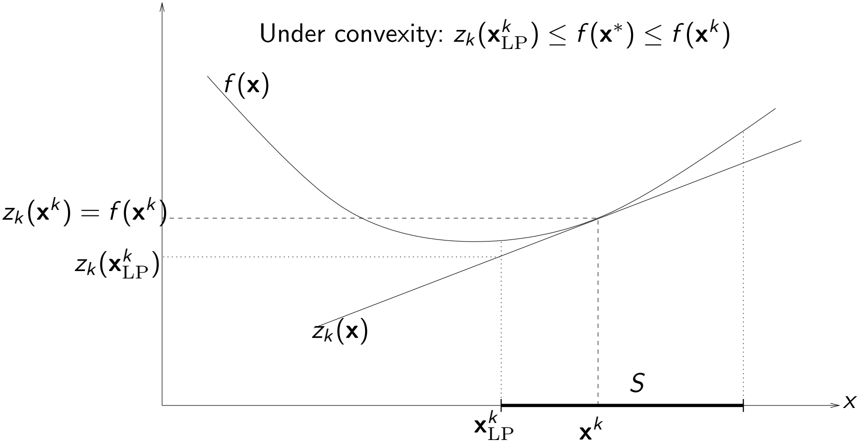
$$\text{minimize } f(\mathbf{x}) \text{ subject to } \mathbf{x} \in S$$

where $f : \Re^n \mapsto \Re$ is convex and $S \subset \Re^n$ is a bounded polyhedron

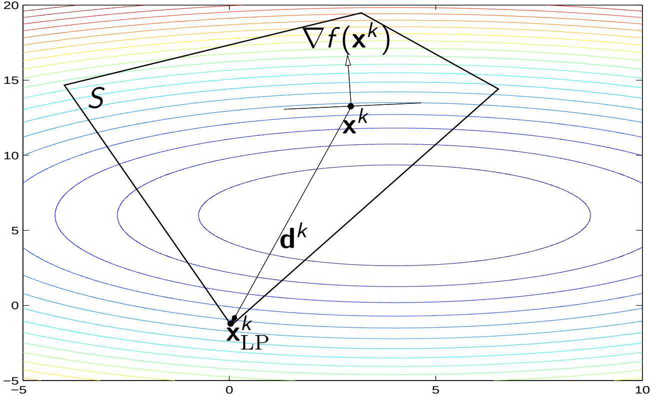
1. Choose $\mathbf{x}^0 \in S$ (simplex, phase one) and $\varepsilon > 0$. Let $\mathtt{UB} = f(\mathbf{x}^0)$, $\mathtt{LB} = -\infty$, $k = 0$
2. Solve the linear approximation (LP):
$$\min_{\mathbf{x} \in S} z_k(\mathbf{x}) := f(\mathbf{x}^k) + \nabla f(\mathbf{x}^k)^T(\mathbf{x} - \mathbf{x}^k) \quad \Rightarrow \quad \mathbf{x} = \mathbf{x}^k_{\mathrm{LP}}$$
Let $\mathbf{d}^k = \mathbf{x}^k_{\mathrm{LP}} - \mathbf{x}^k$ and $\mathtt{LB} = \max\{\mathtt{LB}, z_k(\mathbf{x}^k_{\mathrm{LP}})\}$
If $\mathtt{UB} - \mathtt{LB} < \varepsilon$, stop
3. Solve $\min_{0 \le t \le 1} \varphi(t) := f(\mathbf{x}^k + t \cdot \mathbf{d}^k) \quad \Rightarrow \quad t = t_k$
4. Let $\mathbf{x}^{k+1} = \mathbf{x}^k + t_k \mathbf{d}^k$, $\mathtt{UB} = f(\mathbf{x}^{k+1})$
5. If $\mathtt{UB} - \mathtt{LB} < \varepsilon$, stop. Let $k := k + 1$ and go to step 2

# The Frank–Wolfe-algorithm

- Solves a non-linear optimization problem using
  - a *sequence of approximating*, linear (easier) *problems*, and
  - a sequence of one dimensional (easy) non-linear problems.
- *Estimates* of the optimal objective value is used to terminate the procedure at a *guaranteed maximal deviation* from an optimal solution ($\varepsilon > 0$).



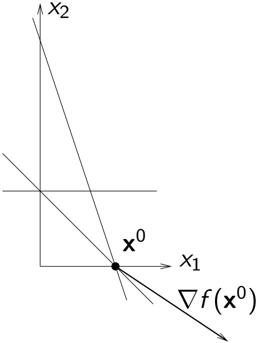

Figur: Illustration of the Frank–Wolfe algorithm in $\Re^1$




Figur: Step 1 of the Frank–Wolfe algorithm.

# An example solved by the Frank–Wolfe-algorithm

$$
\begin{array}{lrcl}
\text{minimize} & f(\mathbf{x}) & = & 3x_1^2 + x_2^2 - x_1x_2 - 3x_2 \\
\text{subject to} & x_1 + x_2 & \geq & 1 \\
 & 3x_1 + x_2 & \leq & 3 \\
 & x_2 & \leq & 1
\end{array}
$$



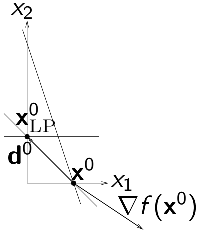

- $\nabla f(\mathbf{x}) = \begin{pmatrix} 6x_1 - x_2 \\ 2x_2 - x_1 - 3 \end{pmatrix}$
- $\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} 6 & -1 \\ -1 & 2 \end{pmatrix}$ positive definite $\Rightarrow$ $f$ strictly convex
- $\mathbf{x}^0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
- $f(\mathbf{x}^0) = 3 \Rightarrow [\mathtt{LB}, \mathtt{UB}] = [-\infty, 3]$

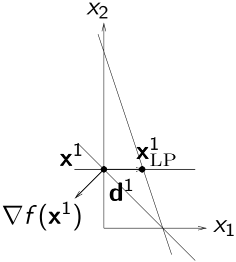
# Frank–Wolfe-example, continued

- $z_0(\mathbf{x}) = 6x_1 - 4x_2 - 3 \Rightarrow \mathbf{x}_{\mathrm{LP}}^0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$
- $z_0(\mathbf{x}_{\mathrm{LP}}^0) = -7 \quad \Rightarrow \quad [\mathrm{LB}, \mathrm{UB}] = [-7, 3]$

$$\left.\begin{array}{l} \mathbf{x}^0 + t \cdot (\mathbf{x}_{\mathrm{LP}}^0 - \mathbf{x}^0) = \begin{pmatrix} 1-t \\ t \end{pmatrix} \\ \varphi(t) = 3(1-t)^2 + t^2 - (1-t)t - 3t \\ \varphi'(t) = 10t - 10 = 0 \quad \Rightarrow \quad t_0 = 1 \end{array}\right\}$$

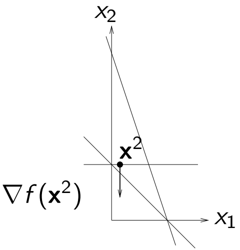
- $\Rightarrow \mathbf{x}^1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$
- $f(\mathbf{x}^1) = -2 \quad \Rightarrow \quad [\mathrm{LB}, \mathrm{UB}] = [-7, -2]$
- $z_1(\mathbf{x}) = -x_1 - x_2 - 1 \Rightarrow \mathbf{x}_{\mathrm{LP}}^1 = \begin{pmatrix} 2/3 \\ 1 \end{pmatrix}$
- $z_1(\mathbf{x}_{\mathrm{LP}}^1) = -\frac{8}{3} \Rightarrow [\mathrm{LB}, \mathrm{UB}] = \left[-\frac{8}{3}, -2\right]$



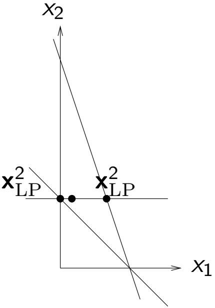



# Frank–Wolfe-example, continued

$$\left.\begin{array}{l} \mathbf{x}^1 + t\cdot(\mathbf{x}^1_{\mathrm{LP}} - \mathbf{x}^1) = \begin{pmatrix} 2t/3 \\ 1 \end{pmatrix} \\ \varphi(t) = 4t^2/3 - 2t/3 - 2 \\ \varphi'(t) = 8t/3 - 2/3 = 0 \Rightarrow t_1 = 1/4 \end{array}\right\}$$




- $\Rightarrow \mathbf{x}^2 = \begin{pmatrix} 1/6 \\ 1 \end{pmatrix}$
- $f(\mathbf{x}^2) = -\frac{25}{12} \Rightarrow [\mathrm{LB}, \mathrm{UB}] = [-\frac{8}{3}, -\frac{25}{12}]$
- $z_2(\mathbf{x}) = -\frac{7}{6}x_2 - \frac{11}{12} \Rightarrow \mathbf{x}^2_{\mathrm{LP}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ or $\begin{pmatrix} 2/3 \\ 1 \end{pmatrix}$
- $z_2(\mathbf{x}^2_{\mathrm{LP}}) = -\frac{25}{12} \Rightarrow [\mathrm{LB}, \mathrm{UB}] = [-\frac{25}{12}, -\frac{25}{12}]$
  Optimal!
- $\mathbf{x}^* = \mathbf{x}^2 = \begin{pmatrix} 1/6 \\ 1 \end{pmatrix}, \quad f(\mathbf{x}^*) = -\frac{25}{12}$



# Penalty function methods (Ch. 12.3)

- Consider both inequality and equality constraints:

$$\begin{aligned} \text{minimize}_{\mathbf{x}\in\Re^n} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & g_i(\mathbf{x}) \le 0, \quad i \in \mathcal{L}, \\ & h_i(\mathbf{x}) = 0, \quad i \in \mathcal{E}. \end{aligned} \tag{1}$$

- Drop the constraints and add terms in the objective that *penalize infeasibile solutions*

$$\text{minimize}_{\mathbf{x}\in\Re^n} \quad F_\mu(\mathbf{x}) := f(\mathbf{x}) + \mu \sum_{i\in\mathcal{L}\cup\mathcal{E}} \alpha_i(\mathbf{x}) \tag{2}$$

where $\mu > 0$ and $\alpha_i(\mathbf{x}) = \begin{cases} = 0 & \text{if } \mathbf{x} \text{ satisfies constraint } i \\ > 0 & \text{otherwise} \end{cases}$

- Common penalty functions (which of these are differentiable?):

$i \in \mathcal{L}$: $\alpha_i(\mathbf{x}) = \max\{0, g_i(\mathbf{x})\}$ or $\alpha_i(\mathbf{x}) = (\max\{0, g_i(\mathbf{x})\})^2$

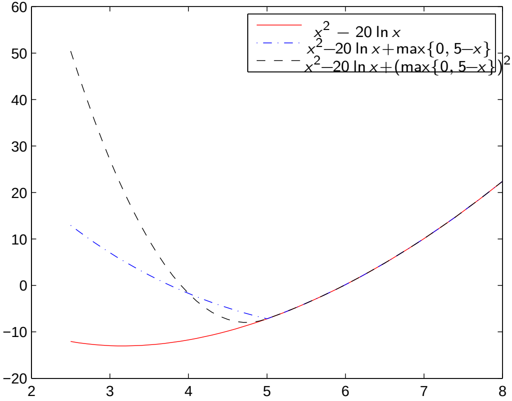
$i \in \mathcal{E}$: $\alpha_i(\mathbf{x}) = |h_i(\mathbf{x})|$ or $\alpha_i(\mathbf{x}) = |h_i(\mathbf{x})|^2$

# More about penalty function methods

- If an optimal solution $\mathbf{x}^*$ to the unconstrained penalty problem (2) is feasible in the original problem (1), it is optimal in (1)
- If the function $g_i$ is differentiable, then the corresponding squared penalty function is also differentiable
- However, squared penalty functions are usually not exact: Typically no value of $\mu > 0$ exists such that an optimal solution for (2) is optimal for the program (1)
- The non-squared penalties are exact: There exists a finite value of $\mu > 0$ such that an optimal solution for (2) is optimal for the program (1)

# Squared and non-squared penalty functions

minimize $x^2 - 20 \ln x$ subject to $x \geq 5$

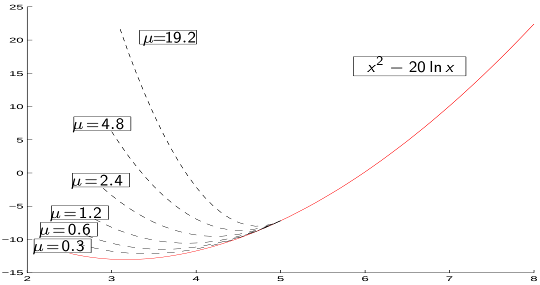



Figur: Squared and non-squared penalty function. $g_i$ differentiable $\Longrightarrow$ squared penalty function differentiable

# Squared penalty functions

- In practice: Start with a low value of $\mu > 0$ and increase the value as the computations proceed
- **Example:** minimize $x^2 - 20 \ln x$ subject to $x \geq 5$ $\quad (*)$

$\Rightarrow$ minimize $x^2 - 20 \ln x + \mu(\max\{0, 5 - x\})^2$ $\quad (**)$

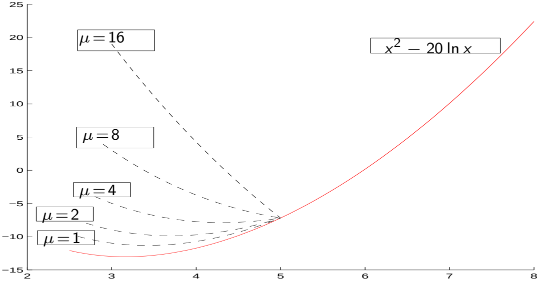



Figur: Squared penalty function: $\nexists \mu < \infty$ such that an optimal solution for $(**)$ is optimal (feasible) for $(*)$

# Non-squared penalty functions

- In practice: Start with a low value of $\mu > 0$ and increase the value as the computations proceed
- **Example:** minimize $x^2 - 20\ln x$ subject to $x \geq 5$ $(+)$

$\Rightarrow$ minimize $x^2 - 20\ln x + \mu \max\{0, 5-x\}$ $(++)$




Figur: Non-squared penalty function: For $\mu \geq 6$ the optimal solution for $(++)$ is optimal (and feasible) for $(+)$

# Sequential unconstrained penalty function algorithm

1. Choose $\mu_0 > 0$, a starting solution $\mathbf{x}^0$, escalation factor $\beta > 1$, and iteration counter $t := 0$
2. Solve

$$\text{minimize}_{\mathbf{x}\in\Re^n} \quad F_\mu(\mathbf{x}) := f(\mathbf{x}) + \mu \sum_{i\in\mathcal{L}\cup\mathcal{E}} \alpha_i(\mathbf{x}) \qquad (2)$$

   with $\mu = \mu_t$, starting from $\mathbf{x}^t \Rightarrow$ optimal solution $\mathbf{x}^{t+1}$
3. If $\mathbf{x}^{t+1}$ is (sufficiently close to) feasible in

$$\begin{aligned} \text{minimize}_{\mathbf{x}\in\Re^n} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & g_i(\mathbf{x}) \le 0, \quad i \in \mathcal{L}, \\ & h_i(\mathbf{x}) = 0, \quad i \in \mathcal{E}. \end{aligned} \qquad (1)$$

   then stop.
   Enlarge the penalty parameter: $\mu_{t+1} := \beta\mu_t$, let $t := t + 1$ and repeat from 2.

# Barrier function methods (Ch. 12.4)

- Consider only inequality constraints:

$$\begin{aligned} \text{minimize}_{\mathbf{x} \in \Re^n} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & g_i(\mathbf{x}) \leq 0, \quad i \in \mathcal{L}. \end{aligned} \tag{3}$$

- Drop the constraints and add terms in the objective that *prevents from approaching the boundary* of the feasible set

$$\text{minimize}_{\mathbf{x} \in \Re^n} \quad F_\mu(\mathbf{x}) := f(\mathbf{x}) + \mu \sum_{i \in \mathcal{L}} \alpha_i(\mathbf{x}) \tag{4}$$

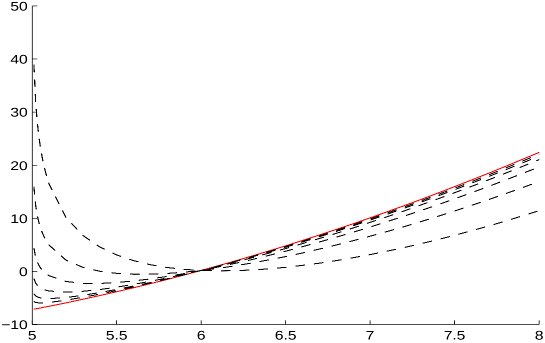
  where $\mu > 0$ and $\alpha_i(\mathbf{x}) \to +\infty$ as $g_i(\mathbf{x}) \to 0$ (as constraint $i$ approaches being active)

- Common barrier functions:
  - $\alpha_i(\mathbf{x}) = -\ln[-g_i(\mathbf{x})]$ or $\alpha_i(\mathbf{x}) = \frac{-1}{g_i(\mathbf{x})}$

# Logarithmic barrier functions

- Choose $\mu > 0$ and decrease it as the computations proceed
- **Example:** minimize $\quad x^2 - 20 \ln x$ subject to $x \geq 5$

$\Rightarrow$ minimize $_{x>5}$ $\; x^2 - 20 \ln x - \mu \ln(x-5)$

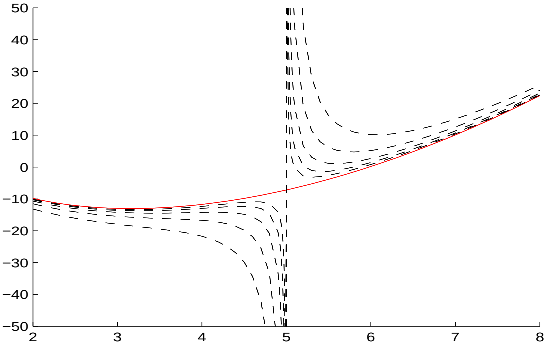



Figur: Logarithmic barrier function: $\mu \in \{10, 5, 2.5, 1.25, 0.625, 0.3125\}$

# Fractional barrier functions

- Choose $\mu > 0$ and decrease it as the computations proceed
- **Example:** minimize $\quad x^2 - 20 \ln x$ subject to $x \geq 5$

$\Rightarrow$ minimize $_{x>5} \quad x^2 - 20 \ln x + \frac{\mu}{x-5}$




Figur: Fractional barrier function: $\mu \in \{10, 5, 2.5, 1.25, 0.625\}$

# More about (fractional) barrier function methods

- If $\mu > 0$ and the true optimum lies on the boundary of the feasible set (i.e., $g_i(\mathbf{x}^*) = 0$ for some $i \in \mathcal{L}$) then the optimum of a barrier function can never equal the true optimum

- Under mild assumptions, the sequence of unconstrained barrier optima converges (in the limit) to the true optimum as $\mu \to 0^+$

# Sequential unconstrained barrier function algorithm

$$\begin{aligned} \text{minimize}_{\mathbf{x} \in \Re^n} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & g_i(\mathbf{x}) \leq 0, \quad i \in \mathcal{L} \end{aligned} \tag{3}$$

1. Choose $\mu_0 > 0$, a feasible interior starting solution $\mathbf{x}^0$ (such that $g_i(\mathbf{x}^0) < 0$, $i \in \mathcal{L}$), reduction factor $\beta < 1$, and iteration counter $t := 0$
2. Solve

$$\text{minimize}_{\mathbf{x} \in \Re^n} \quad F_\mu(\mathbf{x}) := f(\mathbf{x}) + \mu \sum_{i \in \mathcal{L}} \alpha_i(\mathbf{x}) \tag{4}$$

   with $\mu = \mu_t$, starting from $\mathbf{x}^t \Rightarrow$ optimal solution $\mathbf{x}^{t+1}$
3. If $\mu$ is sufficiently small, stop. Otherwise, decrease the barrier parameter: $\mu_{t+1} := \beta\mu_t$, let $t := t + 1$, and repeat from 2.

# Quadratic programming (QP) (Ch. 12.2)

- Example (quadratic convex objective, linear constraints):

$$\begin{array}{lrcl} \text{minimize} & f(\mathbf{x}) & = & -2x_1 - 6x_2 + x_1^2 - 2x_1x_2 + 2x_2^2 \\ \text{subject to} & & & \begin{array}{rcrcl} x_1 & + & x_2 & \leq & 2 \\ -\ x_1 & + & 2x_2 & \leq & 2 \\ x_1 & , & x_2 & \geq & 0 \end{array} \end{array}$$

- General model:

$$\text{minimize } \mathbf{c}^{\mathrm{T}}\mathbf{x} + \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} \ \text{ subject to } \ \mathbf{A}\mathbf{x} - \mathbf{b} \leq \mathbf{0}, -\mathbf{I}\mathbf{x} \leq \mathbf{0}$$

where

$$\mathbf{c} = \begin{pmatrix} -2 \\ -6 \end{pmatrix}, \mathbf{Q} = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}, \mathbf{A} = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 2 \\ 2 \end{pmatrix},$$

$$\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

# QP: The Karush-Kuhn-Tucker conditions

$$\text{minimize } \mathbf{c}^{\mathrm{T}}\mathbf{x} + \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} \quad \text{subject to} \quad \mathbf{A}\mathbf{x} - \mathbf{b} \leq \mathbf{0}, -\mathbf{I}\mathbf{x} \leq \mathbf{0}$$

$$\begin{array}{rcl}
\mathbf{c} + \mathbf{Q}\mathbf{x} + \mathbf{A}^{\mathrm{T}}\boldsymbol{\mu} - \mathbf{I}\boldsymbol{\lambda} & = & 0 \\
\mathbf{A}\mathbf{x} & \leq & \mathbf{b} \\
-\mathbf{I}\mathbf{x} & \leq & \mathbf{0} \\
\boldsymbol{\mu}, \boldsymbol{\lambda} & \geq & \mathbf{0} \\
\boldsymbol{\mu}^{\mathrm{T}}(\mathbf{A}\mathbf{x} - \mathbf{b}) = \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{x} & = & 0
\end{array}$$

Slack variables $\mathbf{s} \geq \mathbf{0}$ of the constraints $\mathbf{A}\mathbf{x} \leq \mathbf{b}$: $\quad \mathbf{A}\mathbf{x} + \mathbf{s} = \mathbf{b}$

$\Rightarrow$ The Karush-Kuhn-Tucker constraints reduce to:

$$\begin{array}{rcl}
\mathbf{Q}\mathbf{x} + \mathbf{A}^{\mathrm{T}}\boldsymbol{\mu} - \mathbf{I}\boldsymbol{\lambda} & = & -\mathbf{c} \\
\mathbf{A}\mathbf{x} + \mathbf{I}\mathbf{s} & = & \mathbf{b} \\
\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \mathbf{s} & \geq & \mathbf{0} \\
\mu_i s_i = \lambda_j x_j & = & 0 \text{ for all } i, j
\end{array}$$

# QP: The Karush-Kuhn-Tucker conditions

- For convex optimization problems $\Rightarrow$ Karush-Kuhn-Tucker conditions are sufficient for a global optimum

$\Rightarrow$ A solution $(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \mathbf{s})$ that fulfils the Karush-Kuhn-Tucker conditions is optimal for convex quadratic programs (QP)

- Not all quadratic programs are convex, though!!!
- The KKT-system is linear, with variables: $\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \mathbf{s} \geq \mathbf{0}$
- Additional conditions: $\mu_i s_i = \lambda_j x_j = 0$ for all $i, j$

$\Rightarrow$ Linear programming: Simplex algorithm with *restricted basis*:

  - Either $\mu_i = 0$ or $s_i = 0$. Either $\lambda_j = 0$ or $x_j = 0$.

  $\Rightarrow$ If, e.g., $s_2$ is in the basis ($s_2 > 0$), $\mu_2$ may *not* enter the basis

  - Introduce artificial variables where needed and solve a phase–1 problem

# QP: The phase–1 problem—The example

- Example (quadratic convex objective, linear constraints):

$$
\begin{array}{lrcl}
\text{minimize} & f(\mathbf{x}) & = & -2x_1 - 6x_2 + x_1^2 - 2x_1x_2 + 2x_2^2 \\
\text{subject to} & & & x_1 + x_2 \leq 2 \\
 & & - & x_1 + 2x_2 \leq 2 \\
 & & & x_1 \;,\; x_2 \geq 0
\end{array}
$$

- 

$$
\begin{array}{llllllllllllll}
\text{minimize} & w = & & & & & & & & & a_1 & +a_2 & & \\
\text{subject to} & 2x_1 & -2x_2 & +\mu_1 & -\mu_2 & -\lambda_1 & & & & +a_1 & & & = & 2 \\
 & -2x_1 & +4x_2 & +\mu_1 & +2\mu_2 & & -\lambda_2 & & & & & +a_2 & = & 6 \\
 & x_1 & +x_2 & & & & & +s_1 & & & & & = & 2 \\
 & -x_1 & +2x_2 & & & & & & +s_2 & & & & = & 2 \\
 & x_1, & x_2, & \mu_1, & \mu_2, & \lambda_1, & \lambda_2, & s_1, & s_2, & a_1, & & a_2 & \geq & 0 \\
 & \multicolumn{13}{l}{\mu_1 s_1 = 0, \quad \mu_2 s_2 = 0, \quad \lambda_1 x_1 = 0, \quad \lambda_2 x_2 = 0}
\end{array}
$$

- Find a starting base by reformulating: $a_1, a_2, s_1, s_2 \Rightarrow$
  $w - a_1 - a_2 = w + 2x_2 + 2\lambda_1 + \lambda_2 - \mu_1 - \mu_2 - 8 = 0$

# The phase–1 problem—reformulated

- Minimize $w$, subject to:

$$
\begin{array}{rrrrrrrrrrrcr}
-w & & -2x_2 & -2\mu_1 & -\mu_2 & +\lambda_1 & +\lambda_2 & & & & & = & -8 \\
 & 2x_1 & -2x_2 & +\mu_1 & -\mu_2 & -\lambda_1 & & & & +a_1 & & = & 2 \\
 & -2x_1 & +4x_2 & +\mu_1 & +2\mu_2 & & -\lambda_2 & & & & +a_2 & = & 6 \\
 & x_1 & +x_2 & & & & & +s_1 & & & & = & 2 \\
 & -x_1 & +2x_2 & & & & & & +s_2 & & & = & 2 \\
 & x_1, & x_2, & \mu_1, & \mu_2, & \lambda_1, & \lambda_2, & s_1, & s_2, & a_1, & a_2 & \geq & 0
\end{array}
$$

under the complementarity conditions:

$$\mu_1 s_1 = \mu_2 s_2 = \lambda_1 x_1 = \lambda_2 x_2 = 0$$

- Solution to the phase–1 problem on next page...

# Solution to the phase–1 problem

| basis | $w$ | $x_1$ | $x_2$ | $\mu_1$ | $\mu_2$ | $\lambda_1$ | $\lambda_2$ | $s_1$ | $s_2$ | $a_1$ | $a_2$ | RHS | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $w$ | -1 | 0 | -2 | -2 | -1 | 1 | 1 | 0 | 0 | 0 | 0 | -8 | $x_2$ in? |
| $a_1$ | 0 | 2 | -2 | 1 | -1 | -1 | 0 | 0 | 0 | 1 | 0 | 2 | $\lambda_2 = 0$ |
| $a_2$ | 0 | -2 | 4 | 1 | 2 | 0 | -1 | 0 | 0 | 0 | 1 | 6 | $\Rightarrow$ OK |
| $s_1$ | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 2 | $s_2$ out |
| $s_2$ | 0 | -1 | **2** | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | |
| $w$ | -1 | -1 | 0 | -2 | -1 | 1 | 1 | 0 | 1 | 0 | 0 | -6 | $\mu_1$ in? |
| $a_1$ | 0 | 1 | 0 | 1 | -1 | -1 | 0 | 0 | 1 | 1 | 0 | 4 | $s_1$ basic |
| $a_2$ | 0 | 0 | 0 | 1 | 2 | 0 | -1 | 0 | -2 | 0 | 1 | 2 | $\Rightarrow$ no |
| $s_1$ | 0 | **3/2** | 0 | 0 | 0 | 0 | 0 | 1 | -1/2 | 0 | 0 | 1 | $x_1$ in? |
| $x_2$ | 0 | -1/2 | 1 | 0 | 0 | 0 | 0 | 0 | 1/2 | 0 | 0 | 1 | OK, $s_1$ out |
| $w$ | -1 | 0 | 0 | -2 | -1 | 1 | 1 | 2/3 | 2/3 | 0 | 0 | -16/3 | $\mu_1$ in? |
| $a_1$ | 0 | 0 | 0 | 1 | -1 | -1 | 0 | -2/3 | 4/3 | 1 | 0 | 10/3 | $s_1 = 0$ |
| $a_2$ | 0 | 0 | 0 | **1** | 2 | 0 | -1 | 0 | -2 | 0 | 1 | 2 | $\Rightarrow$ OK |
| $x_1$ | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 2/3 | -1/3 | 0 | 0 | 2/3 | $a_2$ out |
| $x_2$ | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1/3 | 1/3 | 0 | 0 | 4/3 | |
| $w$ | -1 | 0 | 0 | 0 | 3 | 1 | -1 | 2/3 | -10/3 | 0 | 2 | -4/3 | $s_2$ in? |
| $a_1$ | 0 | 0 | 0 | 0 | -3 | -1 | 1 | -2/3 | **10/3** | 1 | -1 | 4/3 | $\mu_2 = 0$ |
| $\mu_1$ | 0 | 0 | 0 | 1 | 2 | 0 | -1 | 0 | -2 | 0 | 1 | 2 | $\Rightarrow$ OK |
| $x_1$ | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 2/3 | -1/3 | 0 | 0 | 2/3 | $a_1$ out |
| $x_2$ | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1/3 | 1/3 | 0 | 0 | 4/3 | |
| $w$ | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | optimum |
| $s_2$ | 0 | 0 | 0 | 0 | -9/10 | -3/10 | 3/10 | -1/5 | 1 | 3/10 | -3/10 | 2/5 | |
| $\mu_1$ | 0 | 0 | 0 | 1 | 1/5 | -3/5 | -2/5 | -2/5 | 0 | 3/5 | 2/5 | 14/5 | |
| $x_1$ | 0 | 1 | 0 | 0 | -3/10 | -1/10 | 1/10 | 3/5 | 0 | 1/10 | -1/10 | 4/5 | |
| $x_2$ | 0 | 0 | 1 | 0 | 3/10 | 1/10 | -1/10 | 2/5 | 0 | -1/10 | 1/10 | 6/5 | |

# Optimal solution to the phase–1 problem

The optimal solution to the phase–1 problem is given by:

$$\begin{bmatrix} x_1^* = 4/5, & x_2^* = 6/5 \\ \mu_1^* = 14/5, & \mu_2^* = 0 \\ \lambda_1^* = 0, & \lambda_2^* = 0 \\ s_1^* = 0, & s_2^* = 2/5 \end{bmatrix}$$

Note that:
$\mu_1 s_1 = \mu_2 s_2 = \lambda_1 x_1 = \lambda_2 x_2 = 0$

The original QP:

$$\begin{array}{rrcrcrcl} \text{minimize} & f(\mathbf{x}) & = & \multicolumn{5}{l}{-2x_1 - 6x_2 + x_1^2 - 2x_1x_2 + 2x_2^2} \\ \text{subject to} & & & x_1 & + & x_2 & \leq & 2 \\ & & - & x_1 & + & 2x_2 & \leq & 2 \\ & & & x_1 & , & x_2 & \geq & 0 \end{array}$$

$\Rightarrow f(\mathbf{x}^*) = -36/5$

What if $f$ was not convex (i.e., $\mathbf{Q}$ not positive (semi)definite)?

# Graphical illustration