NTNU | Norwegian University of Science and Technology

TTK4135 – Lecture 6
Quadratic Programming
Equality-constrained QPs

Lecturer: Lars Imsland

Purpose of Lecture

- (Very) brief recap LPs

- QPs

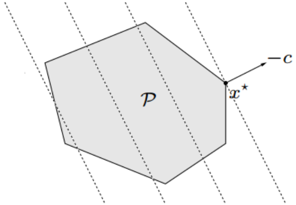
- Solving equality-constrained QPs
  – Can be solved *directly* by solving system of linear equations
  – Two formulations:
    - Full space method: One large equation system
    - Reduced space method: Two smaller equation systems

Reference: Chapter 16.1, 16.2 (15.3) in N&W

Types of Constrained Optimization Problems

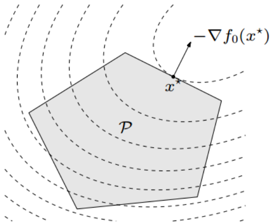
- Linear programming
  - Convex problem
  - Feasible set polyhedron

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



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

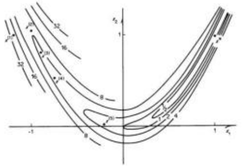
$$\begin{aligned} \min \quad & \frac{1}{2} x^\top P x + q^\top x \\ \text{subject to} \quad & Ax \leq b \\ & Cx = d \end{aligned}$$



- Nonlinear programming
  - In general non-convex!

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

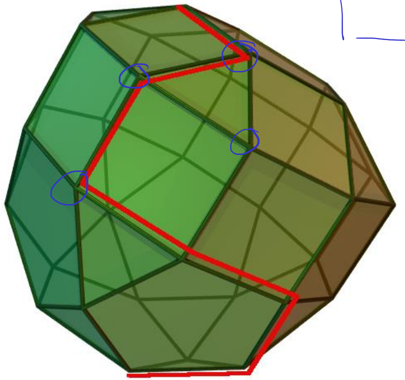
$f(x) = 100\,(x_2 - x_1^2)^2 + (1 - x_1)^2$



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

Last time: The simplex method for LP

$$\min_x \quad c^\top x$$
$$\text{s.t.} \quad Ax = b$$
$$x \geq 0$$



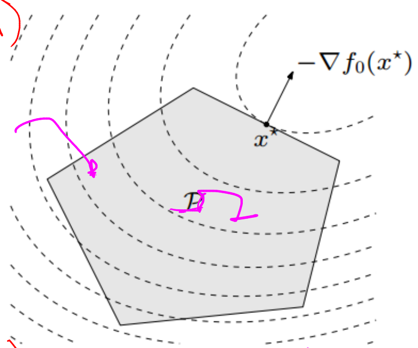
- The Simplex algorithm
  - The feasible set of LPs are (convex) polytopes
  - LP solution is a vertex/“corner”/**BFP** of the feasible set
  - Simplex works by going from vertex to neighbouring vertex in such a manner that the objective decreases in each iteration
  - In each iteration, we solve a linear system to find which component in the **basis** (set of “not active constraints”) we should change
  - “Almost” guaranteed convergence (if LP not unbounded or infeasible)
- Complexity:
  - Typically, at most *2m* to *3m* iterations
  - Worst case: All vertices must be visited (exponential complexity in *n*)
- Active set methods (such as simplex method):
  - Maintains explicitly an estimate of the set of inequality constraints that are active at the solution (the set $\mathcal{N}$ for the Simplex method)
  - Makes small changes to the set in each iteration (a single index in Simplex)

- Next lecture: Active set method for QP

QP: $\min_{x \in \mathbb{R}^n} \underbrace{\frac{1}{2} x^T G x + c^T x}_{q(x)}$, $\quad G = G^T$ (Hessian)

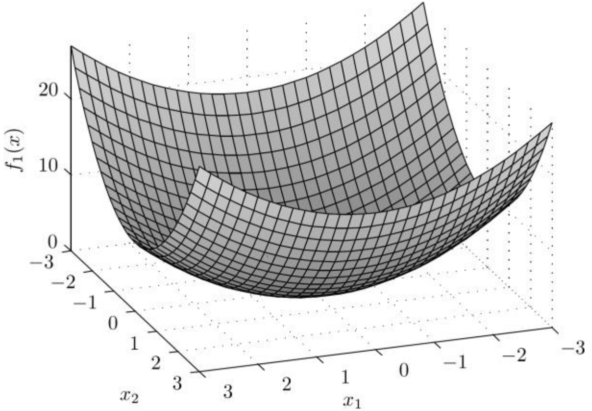
s.t. $\quad c_i(x) = a_i^T x - b_i = 0, \quad i \in \mathcal{E}$

$\quad\quad\; c_i(x) = a_i^T x - b_i \geq 0, \quad i \in \mathcal{I}$

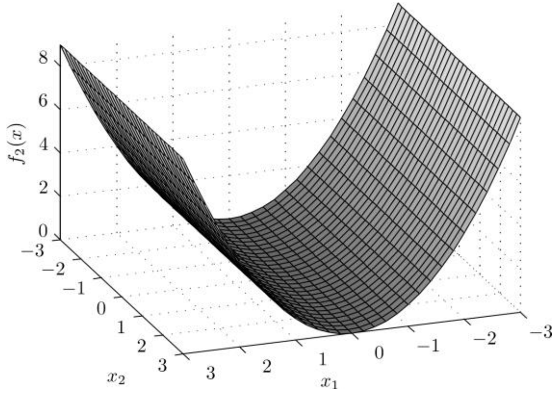


QP is convex if $G \geq 0$ ("Easy case")

convex feasible set



$G > 0$, strictly convex



$G \geq 0$, convex

Why are we interested in (convex) QPs?

- It is the “easiest” nonlinear programming problem
  - “easy”: efficient algorithms exist for convex QPs, when local solutions are global

- The QP is the basic building block of SQP (“sequential quadratic programming”), a common method for solving general nonlinear programs
  - Topic in end of course (N&W Ch. 18)

- QPs are often used in control, especially as solvers in Model Predictive Control
  - Topic in a few weeks
  - Also used in finance (“Portifolio optimization”), some types of Machine Learning/regression problems, control allocation, economics, …

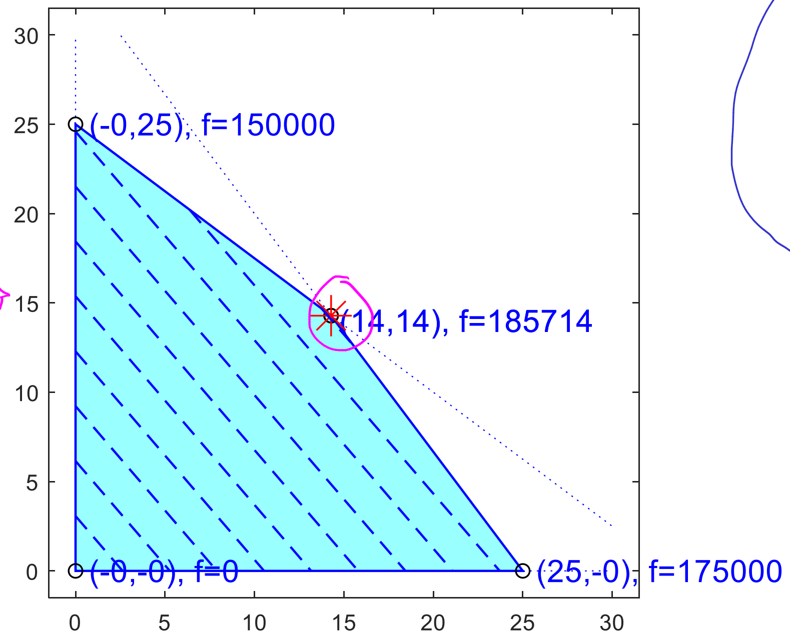
QP Example: Farming example with changing prices

- A farmer wants to grow apples (A) and bananas (B)
- He has a field of size 100 000 m^2
- Growing 1 tonne of A requires an area of 4 000 m^2, growing 1 tonne of B requires an area of 3 000 m^2
- A requires 60 kg fertilizer per tonne grown, B requires 80 kg fertilizer per tonne grown
- The profit for A is 7000 – 200 x_1 per tonne (including fertilizer cost), the profit for B is 6000 – 140 x_1 per tonne (including fertilizer cost)
- The farmer can legally use up to 2000 kg of fertilizer

x_2

- The farmer wants to maximize his profits

LP farming example: Geometric interpretation and solution



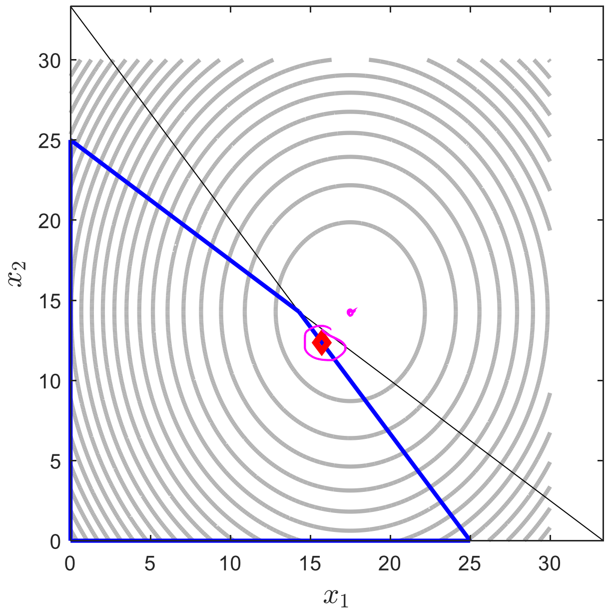
$$
\begin{aligned}
\max_{x_1,\, x_2} \quad & 7000x_1 + 6000x_2 \\
\text{subject to:} \quad & 4000x_1 + 3000x_2 \leq 100000 \\
& 60x_1 + 80x_2 \leq 2000 \\
& x_1 \geq 0 \\
& x_2 \geq 0
\end{aligned}
$$

$$\max_{x_1, x_2} (7000 - 200x_1)x_1 + (6000 - 140x_2)x_2$$

$$\min_{x_1, x_2} 200x_1^2 + 140x_2^2 - 7000x_1 - 6000x_2$$

$$\min_{x \in \mathbb{R}^2} \frac{1}{2} x^T \underbrace{\begin{bmatrix} 400 & 0 \\ 0 & 280 \end{bmatrix}}_{G > 0} x + \underbrace{\begin{bmatrix} -7000 & -6000 \end{bmatrix}}_{c^T} x$$

QP farming example: Geometric interpretation and solution



$$\max_{x_1,\, x_2} \quad (7000 - 200x_1)x_1 + (6000 - 140x_2)x_2$$

$$\begin{aligned} \text{subject to:} \quad 4000x_1 + 3000x_2 &\leq 100000 \\ 60x_1 + 80x_2 &\leq 2000 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

KKT Conditions (Thm 12.1)

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

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

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

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

(complementarity condition/ complementary slackness)

KKT for Equality-constrained QP (EQP)

$b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$

EQP: $\boxed{\begin{aligned} \min_x \quad & \frac{1}{2} x^T G x + c^T x \\ \text{s.t.} \quad & Ax = b \end{aligned}}$

$\mathcal{I} = \emptyset$

$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix} \in \mathbb{R}^{m \times n}, \quad m \leq n$

$\text{rank}(A) = m$

Lagrangian: $\mathcal{L}(x, \lambda) = \frac{1}{2} x^T G x + c^T x - \lambda^T (Ax - b)$

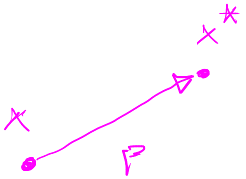
KKT: $\nabla_x \mathcal{L}(x^*, \lambda^*) = G x^* + c - A^T \lambda^* = 0$

$A x^* = b$

Matrix form: $\begin{bmatrix} G & -A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix}$

$x^* = x + p$

KKT for Equality-constrained QP (EQP)



Alternative form: Change variables:

$$x^* = x + p$$

$$\left.\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} -p \\ \lambda^* \end{bmatrix} = \begin{array}{l} \begin{bmatrix} c - Gx \\ Ax - b \end{bmatrix} \end{array}\begin{array}{l} \leftarrow g \\ \leftarrow h \end{array}\right\} \text{KKT system}$$

$\uparrow$ KKT matrix

Observation: When solution exist, easy to solve (LU, LDL)

Questions: - When can it be solved?

- When is the solution a solution to EQP

Solving EQPs

$$\begin{aligned} \min_{x} \quad & \frac{1}{2}x^\top G x + c^\top x \\ \text{subject to} \quad & Ax = b \end{aligned}$$

KKT $\Rightarrow$

$$\begin{pmatrix} G & -A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} x^* \\ \lambda^* \end{pmatrix} = \begin{pmatrix} -c \\ b \end{pmatrix}$$

Solve linear system $\Rightarrow$ $x^*, \quad \lambda^*$

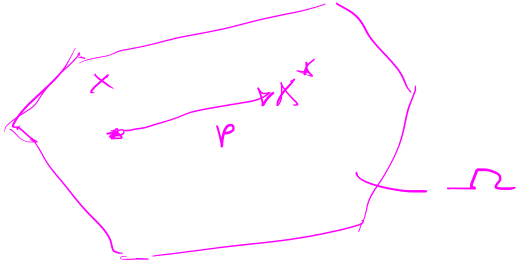
or

$$\begin{pmatrix} G & A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} -p \\ \lambda^* \end{pmatrix} = \begin{pmatrix} c + Gx \\ Ax - b \end{pmatrix}$$

Solve linear system $\Rightarrow$ $x^* = x + p, \quad \lambda^*$

When is the KKT solution the solution to the EQP?

Nullspace



Given x s.t. $Ax=b$.

Rewrite EQP as

$$\min_{p\in\mathbb{R}^n} \frac{1}{2}(x+p)^T G (x+p) + c^T(x+p)$$

$$\text{s.t.} \quad \underbrace{A(x+p)} = b$$

$$Ax + Ap = b \Rightarrow Ap = 0$$

That is: We search for p in $\text{Null}(A) = \{w \mid Aw = 0\}$

Let columns of $Z \in \mathbb{R}^{n\times(m-n)}$ span $\text{Null}(A)$

Ex: $A = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$, $Ap = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = 0$

$A\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0$, $A\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 0$, $Z = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$

When can EQP be solved?

- A full row rank
- $Z^T G Z > 0$

$\xRightarrow{\text{Lemma 16.1}}$ $\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix}$ non-singular

$\Rightarrow \begin{cases} x^* = x + p \\ \lambda^* \end{cases}$ unique solution to KKT system

Thm 16.2

- A full row rank
- $Z^T G Z > 0$

$\Rightarrow$ x^* unique solution to EQP

"Proof" Theorem 16.2

- Assume x feasible, x^* optimal, $x + p = x^*$

$$Ap = A(x^* - x) = Ax^* - Ax = b - b = 0 \Rightarrow p \in \mathrm{Null}(A)$$

- Want to show that $q(x) > q(x^*)$, $x \neq x^*$

$$c^T p = (c^T p)^T = p^T c$$

$$q(x) = \frac{1}{2} x^T G x + c^T x$$

$$= \frac{1}{2}(x^* - p)^T G (x^* - p) + c^T (x^* - p)$$

$$= \underline{\frac{1}{2} x^{*T} G x^*} - \underline{p^T G x^*} + \frac{1}{2} p^T G p + \underline{c^T x^*} - \underline{c^T p}$$

$$= q(x^*) - p^T (\underbrace{G x^* + c}_{\downarrow \text{ KKT}}) + \frac{1}{2} p^T G p \qquad p \in \mathrm{Null}(A)$$

$$\downarrow \qquad p^T A^T \lambda^* = (Ap)^T \lambda^* = 0 \qquad \underline{p = Zu}$$

$$= q(x^*) + \frac{1}{2} u^T \underbrace{Z^T G Z}_{> 0} u > q(x^*)$$

$\square$

Example 16.2

$$\min_x \quad \frac{1}{2}x^\top G x + c^\top x$$
$$\text{subject to} \quad Ax = b$$

$$\min_{x_1,x_2,x_3} \; 3x_1^2 + 2x_1x_2 + x_1x_3 + 2.5x_2^2 + 2x_2x_3 + 2x_3^2 - 8x_1 - 3x_2 - 3x_3$$
$$\text{subject to} \quad x_1 + x_3 = 3, \quad x_2 + x_3 = 0$$

$$\text{Matrices:} \quad G = \begin{pmatrix} 6 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 4 \end{pmatrix}, \quad c = \begin{pmatrix} -8 \\ -3 \\ -3 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$$

Note symmetry of G.
Always possible!

```
>> G = [6 2 1; 2 5 2; 1 2 4]; c = [-8; -3; -3]; A = [1 0 1;0 1 1]; b = [3;0];
>> K = [G, -A'; A, zeros(2,2)];
>> K\[-c;b]     % X = A\B is the solution to the equation A*X = B

ans =

   2.0000
  -1.0000
   1.0000
   3.0000
  -2.0000
```

x^* (2.0000, −1.0000, 1.0000)

λ^* (3.0000, −2.0000)

$$K = \begin{bmatrix} G & -A^T \\ A & 0 \end{bmatrix}$$

$$K \begin{bmatrix} x^* \\ -\lambda^* \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix}$$

Example 16.2

$$\min_{x} \quad \frac{1}{2}x^\top G x + c^\top x$$
$$\text{subject to} \quad Ax = b$$

$$\min_{x_1,x_2,x_3} 3x_1^2 + 2x_1x_2 + x_1x_3 + 2.5x_2^2 + 2x_2x_3 + 2x_3^2 - 8x_1 - 3x_2 - 3x_3$$
$$\text{subject to} \quad x_1 + x_3 = 3, \quad x_2 + x_3 = 0$$

$$\text{Matrices:} \quad G = \begin{pmatrix} 6 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 4 \end{pmatrix}, \quad c = \begin{pmatrix} -8 \\ -3 \\ -3 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$$

$Z^T G Z > 0$

```
>> G = [6 2 1; 2 5 2; 1 2 4]; c = [-8; -3; -3]; A = [1 0 1;0 1 1]; b = [3;0];
>> K = [G, -A'; A, zeros(2,2)];
>> K\[-c;b]     % X = A\B is the solution to the equation A*X = B

ans =

    2.0000
   -1.0000
    1.0000
    3.0000
   -2.0000
```

x^*

λ^*

Solution!

$A = QR$

```
>> [Q,R,P] = qr(A')

Q =

   -0.7071    0.4082   -0.5774
         0   -0.8165   -0.5774
   -0.7071   -0.4082    0.5774

R =

   -1.4142   -0.7071
         0   -1.2247
         0         0

P =

     1     0
     0     1
```

Y Z

$\leftarrow Z$

```
>> Z = Q(:,3);
>> Z'*G*Z

ans =

    4.3333
```

Fundamental Theorem of Linear Algebra

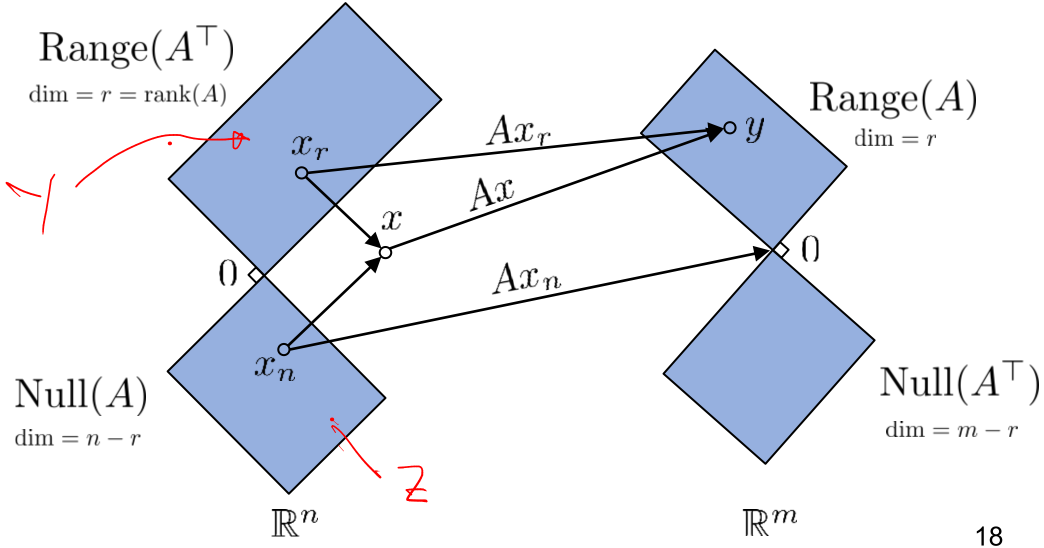
A matrix $A \in \mathbb{R}^{m\times n}$ is a mapping:

Nullspace of A: $\mathrm{Null}(A) = \{v \in \mathbb{R}^n \mid Av = 0\}$

Rangespace (columnspace) of A: $\mathrm{Range}(A) = \{w \in \mathbb{R}^m \mid w = Av, \text{ for some } v \in \mathbb{R}^n\}$

Fundamental theorem of linear algebra:

$$\boxed{\mathrm{Null}(A) \oplus \mathrm{Range}(A^\top) = \mathbb{R}^n}$$



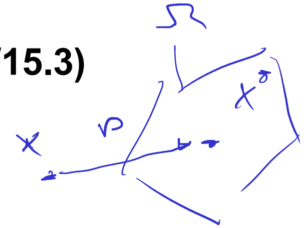
Nullspace method/Elimination of variables (N&W 16.2/15.3)

Given $A \in \mathbb{R}^{m \times n}$, $m \leq n$, $\text{rank}(A) = m$

- Let Z be basis for Null(A), $Z \in \mathbb{R}^{n \times (n-m)}$
- Let Y be basis for Range$(A^\top)$, $y \in \mathbb{R}^{n \times m}$

$$x^* = x + p$$

$$p = \underline{Y} p_Y + \underline{Z} p_Z$$



$$A(x + p) = b$$

$$A(x + Y p_Y + Z p_Z) = b$$

$$Ax + AY p_Y + \underbrace{AZ}_{=0} p_Z = b$$

$$\underbrace{AY}_{\text{non-singular}} p_Y = b - Ax = -h$$

$$\boxed{p_Y = -(AY)^{-1} h.}$$

Note: p_Y completely determined by A, b

Nullspace method/Elimination of variables (N&W 16.2/15.3)

p_Z ? - |

$p = Y p_Y + Z p_Z$

$$q(x+p) = \frac{1}{2}(x+p)^T G (x+p) + c^T (x+p)$$

$$= \underline{\frac{1}{2} x^T G x + c^T x} + \frac{1}{2} p^T G p + \underbrace{(c^T + x^T G)}_{= g^T} p$$

$$= q(x) + \frac{1}{2}(Y p_Y + Z p_Z)^T G (Y p_Y + Z p_Z) + g^T (Y p_Y + Z p_Z)$$

$$= \underbrace{q(x) + \frac{1}{2} p_Y^T Y^T G Y p_Y + g^T Y p_Y}_{\text{known}} + \frac{1}{2} p_Z^T \underbrace{Z^T G Z}_{>0} p_Z + p_Y^T G Z p_Z + g^T Z p_Z$$

$f(p_Z)$

$$\min_p q(x+p) \text{ s.t. } Ap = 0 \Leftrightarrow \min_{p_Z} f(p_Z) \Leftrightarrow \frac{d}{dp_Z} f(p_Z) = 0$$

$$\Leftrightarrow \underbrace{(Z^T G Z) p_Z + Z^T G p_Y + Z^T g = 0}_{\text{linear system in } p_Z}$$

Nullspace method/Elimination of variables (N&W 16.2/15.3)

Summing up: Direct solutions of KKT system (16.2)

Solution of KKT system when $Z^\top G Z > 0$

- Full space:

$$\begin{pmatrix} G & A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} -p \\ \lambda^* \end{pmatrix} = \begin{pmatrix} c + Gx \\ Ax - b \end{pmatrix}$$

  – Use LU (or better, LDL-method, since KKT-matrix is symmetric)

- Reduced space, efficient if *n-m* ≪ *n*:

$$(AY)p_Y = b - Ax \qquad \leftarrow \text{LU}$$

$$(Z^\top GZ)p_Z = -Z^\top GYp_Y - Z^\top(c + Gx) \qquad \leftarrow \text{Cholesky}$$

$$p = Yp_Y + Zp_Z$$

  – Solve two much smaller systems using LU and Cholesky (both with complexity that scales with n^3)
  – Main complexity is calculating basis for nullspace. Usual method is using QR.

- Alternative to direct methods: Iterative methods for linear equation systems (16.3)
  – For very large systems, can be parallelized

Next time

- Active set method for general (convex) QPs
- Solving EQPs are key ingredient in active set method