END332E QUIZ 4 (Morning Section) 08. 12. 2025

Y. İlker Topcu, Özgür Kabak, Zeynep Kızılkan, Hatice Kocaman, Seren Kadıköylü

1) In Affine Scaling Approach move direction is calculated as $\mathbf{\Delta x = \mp Pc}$. Why does the projection matrix $\mathbf{P = I - A^T\left(AA^T\right)^{-1}A}$ generate a feasible direction?

In the affine scaling method, feasibility is defined with respect to the equality constraints

$$Ax = b$$

When we take a step, the new point must also satisfy:

$$A(x + \Delta x) = b$$

This holds only if:

$$A\Delta x = 0$$

$$P = I - A^T(AA^T)^{-1}A$$

$$AP = A(I - A^T(AA^T)^{-1}A) = A - AA^T(AA^T)^{-1}A = A - A = 0$$

$$A\Delta x = A(\mp \mathrm{Pc}) = 0$$

Because of AP=0, any direction calculated using P (such as $\Delta \mathrm{x} = \mp \mathrm{Pc}$) ensures that $\mathrm{A}\Delta\mathrm{x} = 0$. Therefore, the equality constaint remain satisfied, making the direction feasible.

2) In the Affine Scaling Approach, the following equation is used in the final step of an iteration: $\tilde{\boldsymbol{x}}_{\text{new}} = \tilde{\boldsymbol{x}} + \frac{\alpha}{v}\boldsymbol{c}_{\mathrm{P}}$

What do α and v represent in this equation? Explain the rationale for using each of these parameters.

$\boldsymbol{v}$: Negative components of c_P indicate directions in which some variables decrease. If the step is too large, a variable may become zero or negative, violating feasibility ($\tilde{x}_i > 0$). Dividing by v normalizes the step so that the largest decrease is scaled to a safe unit level.
This ensures the algorithm moves in the descent direction without leaving the feasible region.

$\boldsymbol{\alpha}$: controls how much of the maximum feasible step is actually taken.
A larger α gives faster progress toward optimality, while a smaller α improves numerical stability.
α is chosen:

- small enough that we don't approach the bound of the feasible region too much,
- large enough to make progress.

END332E QUIZ 4 (Afternoon Section) 08. 12. 2025

Y. İlker Topcu, Özgür Kabak, Zeynep Kızılkan, Hatice Kocaman, Seren Kadıköylü

Question:
Consider the following non-linear program.

Minimize $f(x_1, x_2, x_3) = x_1 \cdot x_2 + e^{x_2} + (x_3 - 5x_1)^2 - 4x_1$

Subject to

$$x_1^2 + x_2^2 + x_3 \leq 25$$
$$x_1 + 3x_3 \geq 12$$
$$x_2 \geq 0$$

Write down the Karush-Kuhn-Tucker (KKT) conditions necessary for optimality. (Do not solve the system, just list the equations explicitly).

Solution:

$$g_1(x) = {x_1}^2 + {x_2}^2 + x_3 \qquad b_1 = 25$$

$$g_2(x) = x_1 + 3x_3 \qquad b_2 = 12$$

$$g_3(x) = x_2 \qquad b_3 = 0$$

• **complementary slackness conditions**

$$\lambda_1(25 - {x_1}^2 - {x_2}^2 - x_3) = 0$$

$$\lambda_2(12 - x_1 - 3x_3) = 0$$

$$\lambda_3(-x_2) = 0$$

• **sign restrictions**

$$\lambda_1 \leq 0$$

$$\lambda_2 \geq 0$$

$$\lambda_3 \geq 0$$

• **gradient equation**

$$\nabla f(x) = \sum_{i=1}^{3} \lambda_i \nabla g_i(x)$$

$$\nabla f(x) = \begin{pmatrix} x_2 + 2(x_3 - 5x_1)(-5) - 4 \\ x_1 + e^{x_2} \\ 2(x_3 - 5x_1) \end{pmatrix} = \begin{pmatrix} x_2 - 10(x_3 - 5x_1) - 4 \\ x_1 + e^{x_2} \\ 2(x_3 - 5x_1) \end{pmatrix}$$

$$\nabla g_1(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \\ 1 \end{pmatrix}$$

$$\nabla g_2(x) = \begin{pmatrix} -1 \\ 0 \\ -3 \end{pmatrix}$$

$$\nabla g_3(x) = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

$$x_2 - 10(x_3 - 5x_1) - 4 + 2\lambda_1 x_1 - \lambda_2 = 0$$

$$x_1 + e^{x_2} + 2\lambda_1 x_2 - \lambda_3 = 0$$

$$2(x_3 - 5x_1) + \lambda_1 - 3\lambda_2 = 0$$

- **primal constraints of the original NLP model**

$$x_1^2 + x_2^2 + x_3 \leq 25$$

$$x_1 + 3x_3 \geq 12$$

$$x_2 \geq 0$$