1. We have the following optimization problem:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & c(x) \leq 0 \end{aligned} \tag{1}$$

with a dual variable $\lambda$ associated with the inequality constraint $c(x) \leq 0$ and a Lagrangian $\mathcal{L}(x, \lambda) = f(x) + \lambda^T c(x)$. Which of the following is **not** one of the KKT conditions?

(a) $\nabla_x f(x) + \left[\frac{\partial c}{\partial x}\right]^T \lambda = 0$

(b) $c(x) \leq 0$

(c) $\lambda \geq 0$

(d) $c(x)^T \lambda > 0$

(e) $\lambda \odot c(x) = 0$ ($\odot$ is element-wise multiplication)

Solution: (d). (a) is stationarity, (b) is primal feasibility, (c) is dual feasibility, and (e) is complimentarity.

2. If $a \odot b = 0$, does $a^T b = 0$?

(a) yes

(b) no

Solution: (a). If $a \odot b = 0$, that means the elementwise product of $a$ and $b$ is all zeros. The dot product of $a$ and $b$, or $a^T b$, is simply the sum of the elementwise products. This means $a^T b = \sum a \odot b$, and the sum of a zero vector is zero..

3. If we have the following optimization problem:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & c(x) = 0 \end{aligned} \tag{2}$$

with the following KKT conditions:

$$\nabla_x \mathcal{L}(x, \lambda) = \nabla_x f(x) + \left[\frac{\partial c}{\partial x}\right]^T \lambda = 0, \tag{3}$$

$$c(x) = 0. \tag{4}$$

which of the following is the correct way to regularize the Gauss-Newton step computation with a regularizer $\beta > 0$? (The regularizer is shown in red.)

(a) $\begin{bmatrix} \nabla_x^2 f(x) + \beta I & \left[\frac{\partial c}{\partial x}\right]^T \\ \frac{\partial c}{\partial x} & +\beta I \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \end{bmatrix} = \begin{bmatrix} -\nabla_x \mathcal{L}(x, \lambda) \\ -c(x) \end{bmatrix}$

(b) $\begin{bmatrix} \nabla_x^2 f(x) + \beta I & \left[\frac{\partial c}{\partial x}\right]^T \\ \frac{\partial c}{\partial x} & -\beta I \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \end{bmatrix} = \begin{bmatrix} -\nabla_x \mathcal{L}(x, \lambda) \\ -c(x) \end{bmatrix}$

Solution: (b). When we calculate the Jacobian of the above KKT conditions, we get a matrix that is symmetric quasi-definite (more info here `https://vanderbei.princeton.edu/tex/myPapers/sqd6.pdf`). For this case, it means we have the same number of positive eigenvalues as we do primal variables, and the same number of negative eigenvalues as we do dual variables. This means that when we go to regularize this system, we want to push the positive eigenvalues to be a little more positive (hence the $+\beta I$ in the top left block), and we want to push the negative eigenvalues to be more negative (hence the $-\beta I$ in the bottom right block. We can also interpret this from the min-max perspective, where we are minimizing over the primal variable (so these eigenvalues are positive), and maximizing over the dual variable (so these eigenvalues are negative.

4. We have the following optimization problem:

$$\underset{a,b}{\text{minimize}} \quad (a-4)^2 + (2b+3)^2 - 3a + 2b + 3, \tag{5}$$

and we want to solve it in a linear system, where $Ax = b$ with $x = [a, b]^T$. What are $A$ and $b$? To solve this problem, put it in a standard quadratic form and solve for when the gradient equals 0.

(a) $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \quad b = \begin{bmatrix} -7 \\ 4 \end{bmatrix}$

(c) $A = \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}, \quad b = \begin{bmatrix} -11 \\ 8 \end{bmatrix}$

(d) $A = \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}, \quad b = \begin{bmatrix} 11 \\ -14 \end{bmatrix}$

Solution: (d). We can put this cost in quadratic form as the following:

$$(a-4)^2 + (2b+3)^2 - 3a + 2b + 3 = \frac{1}{2}\begin{bmatrix} a \\ b \end{bmatrix}^T \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} -11 \\ 14 \end{bmatrix}^T \begin{bmatrix} a \\ b \end{bmatrix} + 3 \tag{6}$$

which has the following gradient:

$$\nabla_{(a,b)} = \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} -11 \\ 14 \end{bmatrix} \tag{7}$$

When we set this to zero and solve for $a$ and $b$, we get:

$$\begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 11 \\ -14 \end{bmatrix} \tag{8}$$

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}^{-1} \begin{bmatrix} 11 \\ -14 \end{bmatrix} \tag{9}$$

Another way to solve this without putting it in quadratic form is to just take the gradient of the objective function with respect to both a and b, set them both to zero, and solve the system of linear equations.

$$\nabla_a = 2a - 11 = 0 \tag{10}$$

$$\nabla_b = 14 + 8b = 0 \tag{11}$$

which is equivalent to:

$$\begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 11 \\ -14 \end{bmatrix} \tag{12}$$