The Karush-Kuhn-Tucker Theorem

Moritz Kuhn
CDSEM Uni Mannheim

November 2006

1 Problem formulation and the Lagrange function

Consider the following maximization problem

$$\begin{array}{lll} \max\limits_{x\in\mathbb{R}^N} & f(x) & \\ s.t. & g_j(x) \geq 0 & j = 1,\dots,m \\ & h_i(x) = 0 & j = 1,\dots,n \end{array}$$

with $f : \mathbb{R}^N \to \mathbb{R}$, $g_j : \mathbb{R}^N \to \mathbb{R}^p$, $h_i : \mathbb{R}^N \to \mathbb{R}^m$ being continuously differentiable functions.

2 Saddle points of the Lagrangian and Karush-Kuhn-Tucker points

Define the Lagrange function of the problem as

$$L(x, \lambda, \mu) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) + \sum_{i=1}^{n} \mu_i h_i(x)$$

Define a saddle point of the Lagrangian as a tuple

$$(\tilde{x}, \tilde{\lambda}, \tilde{\mu}) \qquad s.t. \qquad L(\tilde{x}, \tilde{\lambda}, \tilde{\mu}) = \min_{\mu,\lambda\geq 0} \max_{x} L(x, \lambda, \mu)$$

We know that

$$\min_{\mu,\lambda\geq 0} L(x, \lambda, \mu) \leq \min_{\mu,\lambda\geq 0} \max_{x} L(x, \lambda, \mu) \leq \max_{x} L(x, \lambda, \mu)$$

i.e. $(\tilde{x}, \tilde{\lambda}, \tilde{\mu})$ is a critical point of $L(x, \lambda, \mu)$, but neither a minimum nor a maximum.

As a next step, we want to establish the connection between a saddle point and the solution to the maximization problem.

Consider the Lagrangian

$$L(x, \lambda, \mu) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) + \sum_{i=1}^{n} \mu_i h_i(x)$$

and the FOC with respect to x, which characterize the critical points of $L(x, \lambda, \mu)$ and are necessary for a maximum

$$\nabla_x f(\tilde{x}) + \sum_{j=1}^{m} \lambda_j \nabla_x g_j(\tilde{x}) + \sum_{i=1}^{n} \mu_i \nabla_x h_i(\tilde{x}) = 0$$

Furthermore, consider the FOC of $L(\tilde{x}, \lambda, \mu)$ with respect to (λ, μ), which are necessary for a minimum of $L(\tilde{x}, \lambda, \mu)$

$$L(\tilde{x}, \lambda, \mu) = f(\tilde{x}) + \sum_{j=1}^{m} \lambda_j g_j(\tilde{x}) + \sum_{i=1}^{n} \mu_i h_i(\tilde{x})$$

Define

$$d(\lambda, \mu) := L(\tilde{x}, \lambda, \mu)$$

The function $d(\lambda, \mu)$ is also called the *dual function* of the problem. Notice that $d(\lambda, \mu)$ is an affine function independent of the functional form of $f(x), g_j(x), h_i(x)$. Since it is a linear programming problem the minimum of the function is either $f(\tilde{x})$ or it does not exist.

$$\min_{\mu, \lambda \geq 0} d(\lambda, \mu) = \begin{cases} f(\tilde{x}) & \text{if } g_j(\tilde{x}) \geq 0 \quad \forall j \quad \textit{and} \quad h_i(\tilde{x}) = 0 \quad \forall i \\ -\infty & \text{else} \end{cases}$$

From this result, we can conclude that every saddle point must be a solution to the original maximization problem. To see why, consider two arguments:

1) A saddle point exists iff $\tilde{x}$ is feasible for the maximization problem, i.e.

$$g_j(\tilde{x}) \geq 0 \quad \forall j \quad \wedge \quad h_i(\tilde{x}) = 0 \quad \forall i$$

(Existence saddle point $\Rightarrow$ feasibility of $\tilde{x}$)

2) The Lagrange function with $(\tilde{\lambda}, \tilde{\mu})$ overestimates the objective function on the interior of the feasible set. To see this, consider the following equivalent problem

$$\max_x f(x) + \sum_{j=1}^{m} I(g_j(x) \geq 0) + \sum_{i=1}^{n} I(h_i(x) = 0)$$

with

$$I(g_j(x) \geq 0) = \begin{cases} 0 & \text{if } g_j(x) \geq 0 \quad \forall j \\ -\infty & \text{else} \end{cases}$$

$$I(h_i(x)) = 0) = \begin{cases} 0 & \text{if } h_i(x) = 0 \quad \forall i \\ -\infty & \text{else} \end{cases}$$

This means we penalize the function for violations of the constraints. This problem is equivalent to the first, if we assume that a solution exists.

In a next step, we replace the "hard" penalty function by "weak" linear penalty functions.

$$\max_x f(x) + \sum_{j=1}^{m} \tilde{\lambda}_j g_j + \sum_{i=1}^{n} \tilde{\mu}_i h_i(x)$$

with

$$\tilde{\lambda}_j \geq 0, \quad \tilde{\mu}_i \geq 0 \quad \forall i, j$$

For feasible values, i.e.

$$g_j(x) \geq 0 \quad and \quad h_i(x) = 0 \quad \forall i, j$$

we clearly overestimate the true objective function. We, therefore, know that there are no other (feasible) choices $\hat{x}$ with a higher value of the objective function than the saddle point of the problem with value $f(\tilde{x})$, i.e.

$$f(\tilde{x}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\mu}) \geq L(\hat{x}, \tilde{\lambda}, \tilde{\mu}) = f(\hat{x}) + \sum_{j=1}^{m} \tilde{\lambda}_j g_j(\hat{x})$$

$$\Leftrightarrow \max_x L(x, \tilde{\lambda}, \tilde{\mu}) \geq L(\hat{x}, \tilde{\lambda}, \tilde{\mu}) \quad \forall \hat{x}$$

Hence, we have shown that every saddle point is a solution to the maximization problem.

$$\text{x is maximizer} \Leftarrow \text{saddle point of } L(x, \lambda, \mu)$$

Furthermore, we have already shown that

$$\begin{gathered}\tilde{\lambda}_j \geq 0 \qquad if \qquad g_j(\tilde{x}) = 0 \\ \tilde{\lambda}_j = 0 \qquad if \qquad g_j(\tilde{x}) > 0 \\ \tilde{\mu}_i \geq 0, \quad h_i(\tilde{x}) \neq 0 \quad \forall i\end{gathered}$$

and

$$\nabla_x f(\tilde{x}) + \sum_{j=1}^{m} \tilde{\lambda}_j \nabla_x g_j(\tilde{x}) + \sum_{i=1}^{n} \tilde{\mu}_i \nabla_x h_i(\tilde{x}) = 0$$

Hence we can conclude that every saddle point satisfies the Karush-Kuhn-Tucker conditions

$$\begin{gathered}\nabla_x f(\tilde{x}) + \sum_{j=1}^{m} \tilde{\lambda}_j \nabla_x g_j(\tilde{x}) + \sum_{i=1}^{n} \tilde{\mu}_i \nabla_x h_i(\tilde{x}) = 0 \\ g_j(\tilde{x}) \geq 0 \quad \forall j \qquad h_i(\tilde{x}) = 0 \quad \forall i \\ \tilde{\lambda}_j \geq 0, \quad \tilde{\mu}_i \geq 0 \quad and \quad \tilde{\lambda}_j g_j(\tilde{x}) = 0 \quad \forall j\end{gathered}$$

Therefore, we get

x is maximizer $\Leftarrow$ saddle point of $L(x, \lambda, \mu) \Rightarrow (\tilde{x}, \tilde{\lambda}, \tilde{\mu})$ satisfy
Karush-Kuhn-Tucker condition

The next question we want to examine is, under which conditions a KKT point, i.e. a point that satisfies the Karush-Kuhn-Tucker conditions, is also a saddle point of the Lagrangian.

As we have shown above, $d(\lambda, \mu)$ is an affine function and we minimize over a convex set $\{(\lambda, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}^n\}$. We, therefore, know that, if a solution exists, then it is $f(\tilde{x})$. Thus, we can conclude that at the saddle point, we have found a minimum with respect to (λ, μ).

The next issue is, under which conditions FOC wrt x are sufficient for a maximum of the Lagrangian. We already know that FOC are sufficient, if we consider a concave programming problem (concave objective function and convex choice set).

Recall the Lagrangian

$$L(x, \lambda, \mu) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) + \sum_{i=1}^{n} \mu_i h_i(x)$$

The Lagrangian is concave if

$$f : \mathbb{R}^N \rightarrow \mathbb{R}, \quad g_j : \mathbb{R} \rightarrow \mathbb{R} \quad \forall j \quad and \quad h_i : \mathbb{R} \rightarrow \mathbb{R} \quad \forall i$$

are concave functions . Remember that the sum of concave functions is concave and that $(\lambda, \mu) \geq 0$.

Now let us check that, with this concavity assumption, the choice set is convex. Assume x' and x'' are feasible, i.e.

$$g_j(x') \geq 0 \quad g_j(x'') \geq 0 \quad \forall j \quad and \quad h_i(x') = 0 \quad h_i(x'') = 0 \quad \forall i$$

What about $\bar{x} = \alpha x' + (1-\alpha)x'' \qquad \alpha \in (0,1)$?

$$\begin{aligned} g_j(\bar{x}) = g_j(\alpha x' + (1-\alpha)x'') \geq \alpha g_j(x') + (1-\alpha)g_j(x'') \geq 0 \quad \forall j \\ h_i(\bar{x}) = h_i(\alpha x' + (1-\alpha)x'') \geq \alpha h_i(x') + (1-\alpha)h_i(x'') \geq 0 \quad \forall i \end{aligned}$$

For feasibility with respect to $h_i(x)$, we see from the second equation that $h_i(\cdot)$ must be affine, i.e. we can write the equality constraints as $Ax = b$, where A is a matrix of dimension $n \times N$ and b is a vector of dimension n.

This yields that

$$h_i(\alpha x' + (1-\alpha)x'') = \alpha h_i(x') + (1-\alpha)h_i(x'') = 0$$

To get sufficient conditions for a maximum of the Lagrangian, we need, therefore, concavity of all functions and that equality constraint can be written as $Ax = b$.

Hence, we update our diagram as follows

$$\begin{gathered} \text{x is maximizer} \Leftarrow \text{saddle point of } L(x,\lambda,\mu) \Rightarrow (\tilde{x}, \tilde{\lambda}, \tilde{\mu}) \text{ satisfy} \\ \text{Karush-Kuhn-Tucker condition} \Leftarrow \text{concavity + h affine} \end{gathered}$$

The last connection to be established is the question of the necessity of a maximizer to be a saddle point of the Lagrangian and, therefore, the necessity of the KKT conditions.

3 Necessity of the Karush-Kuhn-Tucker conditions and the existence of Lagrange multipliers

To get the necessity of the KKT conditions, we have to do a bit more. First, we define a new object. A cone $\mathcal{C}$ in $\mathbb{R}^N$ is the set of points s.t.

$$x \in \mathcal{C} \Rightarrow \lambda x \in \mathcal{C} \qquad \forall \lambda > 0$$

We define furthermore a tangent cone at $\bar{x} \in M$ as

$$\mathcal{T}(\bar{x}, M) = \{d \in \mathbb{R}^N : \exists \alpha_k > 0, \bar{x}_k \in M : \lim_{k \to \infty} \bar{x}_k \to \bar{x} \wedge \lim_{k \to \infty} \alpha_k(\bar{x} - \bar{x}_k) = d\}$$

The d are also called a *limiting direction* of a feasible sequences.

Next, we want to establish the connection between the tangent cone and an optimal solution. To do so we define the set of feasible choices as

$$\mathcal{F} := \{x : g(x) \geq 0 \wedge h(x) = 0\}$$

Lemma 1. *Every solution x^* to the maximization problem satisfies*

$$g(x^*) \geq 0, \quad h(x^*) = 0 \quad \textit{and} \quad (\nabla_x f(x^*))d \geq 0$$
$$\forall d \in \mathcal{T}(x^*, \mathcal{F})$$

Proof. $\alpha_k(f(x^*) - f(x_k)) \geq 0$

by Taylor's theorem, we get

$$\alpha_k(\nabla_x f(x^*)(x^* - x_k) + \alpha_k \wr(\|x^* - x_k\|) \geq 0$$

and for $x_k \to x^*$, we get

$$\nabla_x f(x^*)d \geq 0$$

with

$$d := \lim_{k \to \infty} \alpha_k(x^* - x_k)$$

which proves the lemma.

□

Now define the index set of active constraints

$$A(x) := \{j : g_j(x) = 0\}$$

Using this, we can linearize the tangent cone

$$\mathcal{T}_L(x, g, h) := \{d \in \mathbb{R}^N : \nabla_x g_j(x)d \leq 0 \quad \forall j \in A(x), \nabla_x h(x)d = 0\}$$

Lemma 2. *For continuously differentiable functions*

$$g : \mathbb{R}^N \to \mathbb{R}^m \quad \textit{and} \quad h : \mathbb{R}^N \to \mathbb{R}^n$$

we have

$$\mathcal{T}(x, \{x : g(x) \geq 0, h(x) = 0\}) \subset \mathcal{T}_L(x, g, h)$$

for all feasible x.

Proof. Pick an arbitrary $d \in \mathcal{T}(\bar{x}, \mathcal{F})$ for a feasible $\bar{x}$ together with a feasible sequence $\{x_k\}$ with $x_k \to \bar{x}$ and for $\alpha > 0$, we have

$$\alpha_k(h\bar{x} - h(x_k)) = 0 \Leftrightarrow \alpha_k(\nabla_k h(\bar{x})(\bar{x} - x_k)) + \alpha_k \wr(\|x^* - x_k\|) = 0$$

and for $x_k \to \bar{x}$, we have

$$\nabla_x h(\bar{x})d = 0$$

Furthermore

$$\alpha_k(g_j(\bar{x}) - g_j(x_k)) \leq 0$$

because we only consider $g_j(\bar{x}) = 0$ and $g_j(x_k) \geq 0$

$$\Leftrightarrow \alpha_k \nabla_x g_j(\bar{x})(\bar{x} - x_k) + \alpha_k 0(\|x^* - x_k\|) \leq 0$$

and for $x_k \to \bar{x}$

$$\nabla_x g_j(\bar{x})d \leq 0 \quad \forall j \in A(\bar{x})$$

Hence, lemma 2 is proven.

□

Unfortunately, the reverse is not always true, i.e.

$$\mathcal{T}_L(x, g, h) \not\subset \mathcal{T}(x, \mathcal{F})$$

therefore we impose the following assumption

$$\mathcal{T}_L(x, g, h) = \mathcal{T}(x, \mathcal{F})$$

which is also called the *Abadie constraint qualification* (ACQ) for feasible x. Later on, we consider sufficient conditions for (ACQ). See section 4 about *constraint qualifications.*

Corollary 1. *If x^* is a solution to the maximization problem and (ACQ) holds, then*

$$g(x^*) \geq 0 \quad h(x^*) = 0 \quad \textit{and} \quad \nabla_x f(x^*)d \geq 0 \quad \forall d \in \mathcal{T}_L(x^*, g, h)$$

Now, we need an additional lemma that I will not prove here

Farkas' Lemma 1. *For $A \in \mathbb{R}^{N \times m}, B \in \mathbb{R}^{N \times n}$ and $c \in \mathbb{R}^N$ then*

$$\begin{aligned} & \forall d \in \mathbb{R}^N \quad A^T d \leq 0 \quad B^T d = 0 \quad c^T d \leq 0 \\ \Leftrightarrow \quad & \exists u \in \mathbb{R}^m_+ \quad \textit{and} \quad v \in \mathbb{R}^n \quad \textit{s.t.} \quad An + Bv = c \end{aligned}$$

Using this lemma, we can finally prove that the KKT are necessary for an optimum of the maximization problem.

Kuhn-Tucker Theorem 1. *If x^* is a (local) optimum of the problem*

$$\max_x f(x) \quad \textit{s.t.} \quad g(x) \geq 0 \qquad h(x) = 0$$

and the (ACQ) is satisfied, then

$$\nabla f(x^*) + \nabla g(x^*)\lambda^* + \nabla h(x^*)\mu^* = 0 \tag{1}$$

$$h(x^*) = 0 \qquad g(x^*) \geq 0 \tag{2}$$

$$\lambda^* \geq 0 \qquad \lambda^* g(x^*) = 0 \tag{3}$$

Proof.

$$\nabla g(x^*)\lambda^* + \nabla h(x^*)\mu^* = -\nabla f(x^*) \tag{4}$$

Recall

$$\mathcal{T}_L(x^*, g, h) = \{d \in \mathbb{R}^N : \nabla_x g_j(x) d \leq 0 \quad \forall j \in A(x^*) \ \nabla_x h(x) d = 0\}$$

and from the corollary, we know that

$$\nabla_x f(x^*) d \geq 0 \quad \forall d \in \mathcal{T}_L(x^*, g, h)$$

Since such a d exists, we can use *Farkas' lemma* to conclude that

$$Au + Bv = c \qquad \textit{with} \quad u \geq 0$$

Now replace $A = \nabla_x g(x^*), B = \nabla_x h(x^*)$ and $c = -\nabla_x f(x^*)$ and we get

$$\nabla_x g(x^*) u + \nabla_x h(x^*) v = -\nabla_x f(x^*)$$

If we compare this to (4), we can conclude that there exist $\lambda^* \geq 0$ and μ^* s.t. (4) holds. Hence, we have shown that (1) must hold. ((2)) holds, since x^* is a solution to the optimization problem and ((3)) holds, since $\lambda^* \geq 0$ and $\lambda^*_j = 0$ for $g_j(x^*) > 0$. The theorem is proven.

□

Hence, we can conclude that the KKT conditions are necessary for an optimum. From our previous discussion, we know furthermore that, if f, g are concave and h is affine, then the KKT are sufficient for a saddle path.

Therefore, we can now complete the diagram.

$$x^* \text{ is maximizer} \Leftarrow \text{saddle point of } L(x,\lambda,\mu) \Rightarrow (\tilde{x},\tilde{\lambda},\tilde{\mu}) \text{ satisfy}$$
$$\text{Karush-Kuhn-Tucker condition} \Leftarrow \text{concavity} + \text{h affine}$$

4 Constraint Qualifications

Next we discuss different versions of *constraint qualifications.* A set of conditions is called *constraint qualifications* if the conditions imply that (ACQ) holds. I will not prove that the constraint qualifications are indeed sufficient to imply the (ACQ).

ACQ
We say that the Abadie constraint qualification (ACQ) holds at a feasible point x^* if

$$J(x^*, \mathscr{F}) = J_L(x^*, g, h)$$

where $J(x^*, \mathscr{F})$ denotes the tangent cone of $\mathscr{F}$ at $x^* \in \mathscr{F}$ where $\mathscr{F}$ is the feasible set and $J_L(x^*, g, h)$ denotes the linearized cone at $x^* \in \mathscr{F}$.

LICQ
An easy to check constraint qualification is the linear independence constraint qualification (LICQ). The constraint functions satisfy the (LICQ) iff

$$[\{\nabla_x g_j(x)\}; \{\nabla_x h_i(x)\}]^T \quad \forall j \in \mathscr{A}(x)\ \forall i$$

has full column rank, i.e. the rows of the Jacobian matrices for inequality and equality constraints are linearly independent.
This condition furthermore implies the uniqueless of the Lagrange multipliers.

Corollary 1.
If (LICQ) is satisfied, then the Lagrange multipliers are determined uniquely at a KKT point.

Proof
We know that (x^*, λ^*, μ^*) must satisfy

$$\nabla f(x^*) + \sum_{j=1}^{m} \lambda_j^* \nabla_x g_j(x^*) + \sum_{i=1}^{n} \mu_i^* \nabla_x h_i(x^*) = 0$$

to be a KKT point.

$$\begin{array}{rcl} & \Leftrightarrow & \\ -\nabla f(x^*) & = & \sum_{j \in \mathscr{A}} \lambda_j^* \nabla_x g_j(x^*) + \sum_{i=1}^{m} \mu_i^* \nabla_x h_i(x^*) \\ -\nabla f(x^*) & = & \left[\tilde{\lambda}^*, \mu^*\right]^T \left[\begin{array}{c} D_x \tilde{g}(x^*) \\ D_x h(x^*) \end{array}\right] \end{array}$$

where $\tilde{\lambda}^* = \{\lambda_j^*\}\ j \in \mathscr{A}(x)$, $\tilde{g}(x^*) = \{g_j(x^*\})\ j \in \mathscr{A}(x)$ since $[D_x \tilde{g}(x^*), D_x h(x^*)]^T$ are linearly independent by (LICQ) $[\tilde{\lambda^*}, \mu^*]$ are uniquely determined.

MFCQ
We say that the Mangasarian-Fromovitz constraint qualification (MFCQ) holds at a feasible point x^* if the gradient vectors

$$\nabla h_i(x^*) = 0 \quad \forall i$$

are linearly independent and that $\exists d \in \mathbb{R}^N$ s.t.

$$\begin{array}{rcl} \nabla g_j(x^*)d & < & 0 \ \ \forall j \in \mathscr{A}(x^*) \\ \nabla h_i(x^*)d & = & 0 \ \ \forall i \end{array}$$

It can be shown that (LICQ) ⇒ (MFCQ) and therefore we can conclude that (LICQ) ⇒ (MFCQ) ⇒ (ACQ).

5 Further Constraint Qualifications

Definition 1 (CQ I:).
All constraints $g_j(x) \geq 0$*,* $h_i(x) = 0$ *are all affine.*

Definition 2 (CQ II).
All $h_i(x) = 0$ *are affine and* $g_j(x) \geq 0$ *are convex.*

Remark 1.
Although this seems to imply a well-behaved concave program you show be aware that it actually does not, because the choice set is not convex for this problem.

Definition 3 (Slater's condition)**.**
$h(x)$ *is affine and* $g(x)$ *is concave and* $\exists \bar{x} > 0$*, i.e. there exists a strictly point, and we have a concave programming problem.*

Example

$$\begin{aligned} \max f(x,y) &= x^2 + x + 4y^2 \\ s.t. \qquad & 2x + 2y \leq 1 \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

Since all constraints are linear we know that the constraint qualification and hence (ACQ) is satisfied Hessian of the objective function

$$\mathscr{H} = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}$$

obviously the objective function is convex.
Form the Lagrangian of the problem

$$2(x, y, \lambda, \mu_1, \mu_2) = x^2 + x + 4y^2 + \lambda(1 - 2x - 2y) + \mu_1 x + \mu_2 y$$

we know that if (ACQ) holds that KKT are necessary

$$\begin{aligned} (1) \qquad & 2x + 1 - 2\lambda + \mu_1 = 0 \\ (2) \qquad & 8y - 2\lambda + \mu_2 = 0 \\ (3) \qquad & 1 - 2x - 2y \geq 0 \\ (4) \qquad & \lambda \geq 0 \\ (5) \qquad & \mu_1 \geq 0 \\ (6) \qquad & \mu_2 \geq 0 \\ (7) \qquad & \lambda(1 - 2x - 2y) = 0 \\ (8) \qquad & \mu_1 x = 0 \\ (9) \qquad & \mu_2 y = 0 \\ (10) \qquad & x \geq 0 \\ (11) \qquad & y \geq 0 \end{aligned}$$

from (1) $\quad 2x + 1 = 2\lambda - \mu_1$
if $x = 0 :\ 1 = 2\lambda - \mu_1 \Rightarrow 1 + \mu_1 = 2\lambda \Rightarrow 2x + 2y + 1$
if $x > 0 :\ 2x + 1 = \lambda \Rightarrow 2x + 2y = 1$

from (2) $\quad 8y = 2\lambda - \mu_2$
if $y = 0 :\ \mu_2 = 2\lambda \quad \mu_2 > 0 \Rightarrow 2x + 2y = 1$
if $y > 0 :\ 8y = 2\lambda \Rightarrow 2x = 2y + 1$

Therefore we know that in any case the first constraint must be binding and we have

$$x + y = \frac{1}{2}$$

CASE 1:
$y = 0 \;\; x = \frac{1}{2} \Rightarrow \mu_2 > 0 \;\; \mu_1 = 0$

from (1)

$$\begin{aligned} 1 + 1 - 2\lambda &= 0 \\ \Rightarrow \lambda &= 1 \end{aligned}$$

from (2)

$$\begin{aligned} -2 + \mu_2 &= 0 \\ \Rightarrow \mu_2 &= 2 \end{aligned}$$

CASE 2:
$y = \frac{1}{2} \;\; x = 0 \Rightarrow \mu_2 = 0 \;\; \mu_1 > 0$

from (2)

$$\begin{aligned} 4 - 2\lambda &= 0 \\ \Rightarrow \lambda &= 2 \end{aligned}$$

from (1)

$$\begin{aligned} 1 - 4 + \mu_1 &= 0 \\ \Rightarrow \mu_1 &= 3 \end{aligned}$$

CASE 3:
$x + y = \frac{1}{2} \;\; x > 0 \;\; y > 0 \;\; \mu_1 = 0 \;\; \mu_2 = 0$

$$\begin{aligned} &\text{from (1):} \quad 2x + 1 = 2\lambda \\ &\text{from (2):} \quad 8y = 2\lambda \end{aligned}$$

$$\Leftrightarrow$$

$$\begin{aligned}
2x + 1 &= 8y \\
2\left(\frac{1}{2} - y\right) + 1 &= 8y \\
2 &= 10y \\
y &= \frac{1}{5} \\
\Rightarrow \quad x &= \frac{3}{10} \\
\Rightarrow \quad \lambda &= \frac{4}{5}
\end{aligned}$$

Now evaluate the objective function at the candidates

$$\begin{aligned}
x = \frac{1}{2} \quad y = 0 : &\qquad \frac{1}{4} + \frac{1}{2} = \frac{3}{4} \\
x = 0 \quad y = \frac{1}{2} : &\qquad \frac{1}{4}4 = 1 \quad \longleftarrow \quad \textbf{\textit{optimal solution}} \\
x = \frac{3}{10} \quad y = \frac{1}{5} : &\qquad \frac{9}{100} + \frac{3}{10} + 4\frac{1}{25} = \frac{9}{100} + \frac{30}{100} + \frac{16}{100} = \frac{55}{100} = \frac{11}{22}
\end{aligned}$$