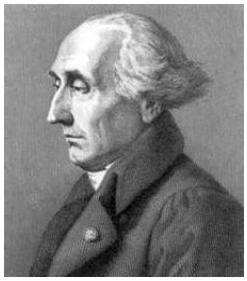
# Chapter 11
## Optimization with Equality Constraints

Albert William Tucker (1905-1995)



Joseph-Louis (Giuseppe Lodovico), comte de Lagrange (1736-1813)

Harold William Kuhn (1925)

1

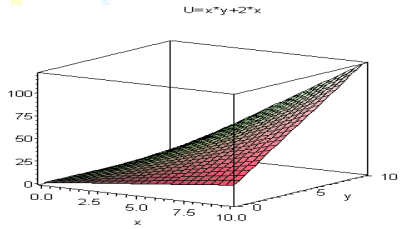
## 11.1 General Problem

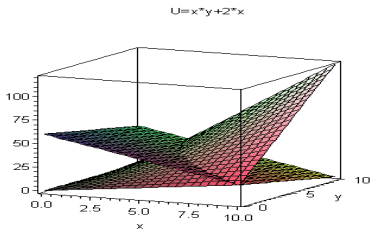
- Now, a constraint is added to the optimization problem:

$$\max_{x,y} u(x, y) \qquad \text{s.t } x\, p_x + y\, p_y = I,$$

$p_x$, $p_y$, & I are exogenous prices and income, respectively.

- Different methods to solve this problem:
  - Substitution
  - Total differential approach
  - Lagrange Multiplier




2

## 11.1 Substitution Approach

- Easy to use in simple 2x2 systems. Using the constraint, substitute into objective function and optimize as usual.
- **Example:**

$U = U(x_1, x_2) = x_1 x_2 + 2x_1$ s.t. $60 = 4x_1 + 2x_2$

1) Solve for $x_2$

$x_2 = 30 - 2x_1$

2) Substituting into $U(x_1, x_2)$

$U = x_1(30 - 2x_1) + 2x_1 = 32x_1 - 2x_1^2$

3) F.o.c.:

$dU/dx_1 = 32 - 4x_1 = 0; \quad \Rightarrow x_1^* = 8; \text{ and } x_2^* = 14;$

Check s.o.c.:

$d^2U/dx_1^2 = -4 < 0 \Rightarrow \text{maximum}$

4) Calculate maximum Value for U(.): $U^* = 128$

3

## 11.2 Total Differential Approach

- Total differentiation of objective function and constraints:

1-2) $U = f(x, y)$; s.t. $B = g(x, y)$

3–4) $dU = f_x dx + f_y dy = 0$; $dB = g_x dx + g_y dy = 0$

5–6) $dU = \begin{bmatrix} f_x & 0 \\ 0 & f_y \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix}$; $dB = \begin{bmatrix} g_x & 0 \\ 0 & g_y \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix}$

7–8) $dx/dy = -f_y/f_x$; $dx/dy = -g_y/g_x$

9–10) $f_y/f_x = g_y/g_x$; $f_y/g_y = f_x/g_x$

- Equation (10) along with the restriction (2) form the basis to solve this optimization problem.

4

## 11.2 Total Differential Approach

- **Example**: $U(x_1, x_2) = x_1x_2 + 2x_1$ s.t. $60 = 4x_1 + 2x_2$

  Taking first-order differentials of U and budget constraint (B):

$$dU = x_2dx_1 + x_1dx_2 + 2dx_1 = 0; \qquad dB = -4dx_1 - 2dx_2 = 0$$

$$dx_1/dx_2 = -x_1/(x_2+2); \qquad dx_1/dx_2 = -1/2$$

$$-x_1/(x_2+2) = -1/2 \qquad \Rightarrow x_1 = (x_2+2)/2$$

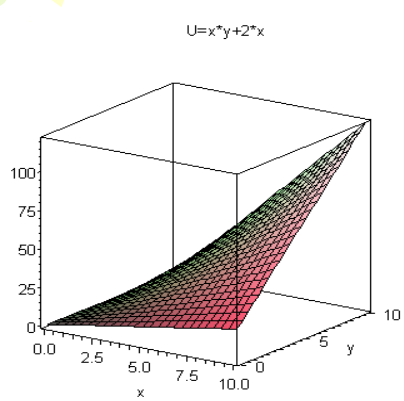
$$60 = 4(1+1/2\,x_2) + 2x_2 \qquad \Rightarrow x_2^* = 14;\ x_1^* = 8$$

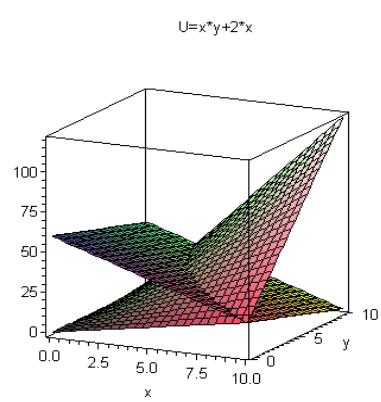
$$U^*(8,14) = (8)14 + 2(8) = 128$$

5

## 11.2 Total-differential approach

- Graph for Utility function and budget constraint:




6

## 11.3 Lagrange-multiplier Approach

- To avoid working with (possibly) zero denominators, let λ denote the common value in (10). Rewriting (10) and adding the budget constraint we are left with a 3x3 system of equations:

$$
\begin{aligned}
&10'') \qquad f_x = \lambda g_x \\
&10'') \qquad f_y = \lambda g_y \\
&2) \qquad B = g(x, y)
\end{aligned}
$$

- There is a convenient function that produces (10'), (10") and (2) as a set of f.o.c.: The *Lagrangian function*, which includes the objective function and the constraint:

$$L = f(x_1, x_2) + \lambda[B - g(x_1, x_2)]$$

- The constraint is multiplied by a variable, λ, called the Lagrange multiplier (LM).

7

## 11.3 Lagrange-multiplier Approach

- Once we form the Lagrangian function, the Lagrange function becomes the new objective function.

$$
\begin{aligned}
&(1) \qquad L = f(x_1, x_2) + \lambda[B - g(x_1, x_2)] \\
&(2) \qquad L_\lambda = B - g(x_1, x_2) = 0 \\
&(3) \qquad L_{x_1} = f_{x_1} - \lambda g_{x_1} = 0 \\
&(4) \qquad L_{x_2} = f_{x_2} - \lambda g_{x_2} = 0 \\
&(5) \qquad H = \begin{bmatrix} L_{\lambda\lambda} & L_{\lambda x_1} & L_{\lambda x_2} \\ L_{x_1\lambda} & L_{x_1 x_1} & L_{x_1 x_2} \\ L_{x_2\lambda} & L_{x_2 x_1} & L_{x_2 x_2} \end{bmatrix}
\end{aligned}
$$

8

## 11.3 LM Approach

- Note that

$$L_{\lambda\lambda} = 0$$
$$L_{\lambda x1} = -g_{x1}$$
$$L_{\lambda x2} = -g_{x2}$$

- Then

$$(5) \qquad H = \begin{bmatrix} 0 & -g_{x_1} & -g_{x_2} \\ -g_{x_1} & L_{x_1x_1} & L_{x_1x_2} \\ -g_{x_2} & L_{x_2x_1} & L_{x_2x_2} \end{bmatrix}$$

- If the constraints are linear, the Hessian of the Lagrangian can be seen as the Hessian of the original objective function, bordered by the first derivatives of the constraints. This new Hessian is called *bordered Hessian.*

9

## 11.3 LM Approach: Example

Maximize Utility $U = U(x,y)$ where $U_x, U_y > 0$

Subject to the budget constraint $B = xP_x + yP_y$

$$L = U(x,y) + \lambda(B - xP_x - yP_y)$$

$$L_\lambda = \beta - xP_x - yP_y = 0$$
$$L_x = U_x - \lambda P_x = 0$$
$$L_y = U_y - \lambda P_y = 0$$

$$L_B = \lambda = \frac{U_x}{P_x} = \frac{U_y}{P_y}$$

$$|H| = \begin{vmatrix} L_{\lambda\lambda} & L_{\lambda x} & L_{\lambda y} \\ L_{x\lambda} & L_{xx} & L_{xy} \\ L_{y\lambda} & L_{yx} & L_{yy} \end{vmatrix} = \begin{vmatrix} 0 & -P_x & -P_y \\ -P_x & U_{xx} & U_{xy} \\ -P_y & U_{yx} & U_{yy} \end{vmatrix}$$

$$= 2P_xP_yU_{xy} - P_y^2U_{xx} - P_x^2U_{yy}$$

10

## 11.3 LM Approach: Second-order conditions

- $\lambda$ has no effect on the value of $L^*$ because the constraint equals zero but …
- A new set of second-order conditions are needed
- The constraint changes the criterion for a relative max. or min.

(1) $H = ax^2 + 2hxy + by^2 \quad s.t. \quad \alpha x + \beta y = 0$

(2) $y = -\frac{\alpha}{\beta}x$ solve the constraint for y

(3) $H = ax^2 + 2hx\left(-\frac{\alpha}{\beta}x\right) + b\left(-\frac{\alpha}{\beta}x\right)^2$

(4) $H = \left(a\beta^2 - 2\alpha\beta h + b\alpha^2\right)\left(\frac{x}{\beta}\right)^2$

(5) $H > 0$ iff $a\beta^2 - 2\alpha\beta h + b\alpha^2 > 0$

11

## 11.3 LM Approach: Second-order conditions

(1) $H = a\beta^2 - 2\alpha\beta h + b\alpha^2 > 0$

(2) $H = \begin{vmatrix} 0 & \alpha & \beta \\ \alpha & a & h \\ \beta & h & b \end{vmatrix} = -a\beta^2 + 2\alpha\beta h - b\alpha^2$

(3) $H$ is positive definite s.t. $\alpha x + \beta y = 0$

iff $\begin{vmatrix} 0 & \alpha & \beta \\ \alpha & a & h \\ \beta & h & b \end{vmatrix} < 0 \quad \rightarrow \min$

12

## 11.3 LM Approach: Example

1-2) $U = x_1x_2 + 2x_1$ s.t. $B = 60 - 4x_1 - 2x_2 = 0$

Form the Lagrangian function

3) $L = x_1x_2 + 2x_1 + \lambda(60 - 4x_1 - 2x_2)$

*FOC*

4) $L_\lambda = 60 - 4x_1 - 2x_2 = 0$

5 – 6) $L_{x_1} = x_2 + 2 - \lambda 4 = 0;$ $\lambda = (1/4)x_2 + 1/2$

7 – 8) $L_{x_2} = x_1 - \lambda 2 = 0;$ $\lambda = (1/2)x_1$

9 – 10) $(1/4)x_2 + 1/2 = (1/2)x_1;$ $x_2 = 2x_1 - 2$

11 – 12) $60 = 4x_1 + 2(2x_1 - 2);$ $x_1^* = 8$

13 – 14) $60 = 4(8) - 2x_2;$ $x_2^* = 14$

15 – 17) $U = (8)(14) + 2(8);$ $U^* = 128;$ $\lambda^* = 4$

13

## 11.3 LM Approach: Restricted Least Squares

- The Lagrangean approach

$$Min_{\beta,\lambda} \quad L(\beta, \lambda \mid y, x) = \sum_{t=1}^{T} (y_t - x_t\beta)^2 + 2\lambda(r\beta - q)$$

f.o.c:

$$\frac{\partial L}{\partial \beta} = \sum_{t=1}^{T} 2(y_t - x_t b^*)(-x_t) + 2\lambda r = 0 \qquad \Rightarrow -\sum_{t=1}^{T} (y_t x_t - x_t^2 b^*) + \lambda r = 0$$

$$\frac{\partial L}{\partial \lambda} = 2(rb^* - q) = 0 \qquad \Rightarrow (rb^* - q) = 0$$

- Then, from the 1st equation:

$$-(x'y - x'xb^*) + \lambda r = 0 \qquad \Rightarrow b^* = (x'x)^{-1}x'y - (x'x)^{-1}\lambda r$$

$$= b - (x'x)^{-1}\lambda r$$

14

## 11.3 LM Approach: Restricted Least Squares

- $\mathbf{b}^* = \mathbf{b} - \mathbf{r}\,(\mathbf{x}'\mathbf{x})^{-1}\boldsymbol{\lambda}$

- Premultiply both sides by **r** and then subtract **q**

$$\mathbf{r}\mathbf{b}^* - \mathbf{q} = \mathbf{r}\mathbf{b} - \mathbf{r}^2\,(\mathbf{x}'\mathbf{x})^{-1}\boldsymbol{\lambda} - \mathbf{q}$$

$$\mathbf{0} = -\,\mathbf{r}^2\,(\mathbf{x}'\mathbf{x})^{-1}\boldsymbol{\lambda} + (\mathbf{r}\mathbf{b} - \mathbf{q})$$

Solving for $\boldsymbol{\lambda}$ $\Longrightarrow$ $\boldsymbol{\lambda} = [\mathbf{r}^2\,(\mathbf{x}'\mathbf{x})^{-1}]^{-1}\,(\mathbf{r}\mathbf{b} - \mathbf{q})$

Substituting in $\mathbf{b}^*$ $\Longrightarrow$ $\mathbf{b}^* = \mathbf{b} - (\mathbf{x}'\mathbf{x})^{-1}\mathbf{r}\,[\mathbf{r}^2\,(\mathbf{x}'\mathbf{x})^{-1}]^{-1}\,(\mathbf{r}\mathbf{b} - \mathbf{q})$

15

## 11.3 LM Approach: N-variable Case

a) No one-variable test because there must be one more variable than constraint

b) 2 – variable test of soc

| | | |
|---|---|---|
| $\lvert\overline{H}_2\rvert > 0$ | negative definite :( | max |
| $\lvert\overline{H}_2\rvert < 0$ | positive definite :) | min |

c) 3 – variable test of soc

| | | |
|---|---|---|
| $\lvert\overline{H}_2\rvert > 0,\quad \lvert\overline{H}_3\rvert < 0$ | negative definite :( | max |
| $\lvert\overline{H}_2\rvert < 0,\quad \lvert\overline{H}_3\rvert < 0$ | positive definite :) | min |

d) n – variable case soc, (p. 361)

| | | |
|---|---|---|
| $\lvert\overline{H}_2\rvert > 0, \lvert\overline{H}_3\rvert < 0, \lvert\overline{H}_4\rvert > 0, \ldots (-1)^n \lvert\overline{H}_n\rvert > 0$ | negative definite :( | max |
| $\lvert\overline{H}_2\rvert < 0, \lvert\overline{H}_3\rvert < 0, \lvert\overline{H}_4\rvert < 0, \ldots, \lvert\overline{H}_n\rvert < 0$ | positive definite :) | min |

Where

$$\lvert\overline{H}_2\rvert = \begin{bmatrix} 0 & g_1 & g_2 \\ g_1 & Z_{11} & Z_{12} \\ g_2 & Z_{21} & Z_{22} \end{bmatrix}; \quad \lvert\overline{H}_3\rvert = \begin{bmatrix} 0 & g_1 & g_2 & g_3 \\ g_1 & Z_{11} & \cdots & \cdots \\ g_2 & \cdots & Z_{22} & \cdots \\ g_3 & \cdots & \cdots & Z_{33} \end{bmatrix};$$

16

## 11.4 Optimality Conditions – Unconstrained Case

- Let $\mathbf{x}^*$ be the point that we think is the minimum for $f(\mathbf{x})$.
- *Necessary condition (for optimality):*

  $df(\mathbf{x}^*) = 0$
- A point that satisfies the necessary condition is a stationary point
  - It can be a minimum, maximum, or saddle point

- Q: How do we know that we have a minimum?
- *Answer: Sufficiency Condition*:

  The sufficient conditions for $\mathbf{x}^*$ to be a strict local minimum are:

  $df(\mathbf{x}^*) = 0$

  $d^2f(\mathbf{x}^*)$ is positive definite

## 11.4 Constrained Case – KKT Conditions

- To prove a claim of optimality in constrained minimization (or maximization), we have to check the found point ($\mathbf{x}^*$) with respect to the (Karesh) Kuhn Tucker (KKT) conditions.

- Kuhn and Tucker extended the Lagrangian theory to include the general classical single-objective nonlinear programming problem:

| | | |
|---|---|---|
| minimize | $f(\mathbf{x})$ | |
| Subject to | $g_j(\mathbf{x}) \le 0$ | for $j = 1, 2, ..., M$ |
| | $h_k(\mathbf{x}) = 0$ | for $k = 1, 2, ..., K$ |
| | $\mathbf{x} = (x_1, x_2, ..., x_N)$ | |

<u>Note</u>: M inequality constraints, K equality constraints.

## 11.4 Interior versus Exterior Solutions

• *Interior*: If constraints are ***inactive*** and (thus) the solution lies at the interior of the feasible space, then the necessary condition for optimality is same as for unconstrained case:

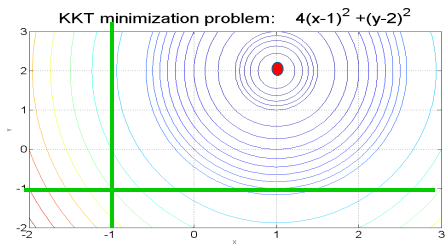
$\nabla f(\mathbf{x}^*) = 0$ ($\nabla$ difference operator for matrices --"del" )

**Example**: Minimize

$f(x) = 4(x-1)^2 + (y-2)^2$

with constraints:

$x \geq -1$ & $y \geq -1$.




*Exterior*: If solution lies at the exterior, the condition $\nabla f(\mathbf{x}^*) = 0$ does *not* apply because some constraints will block movement to this minimum.

## 11.4 Interior versus Exterior Solutions

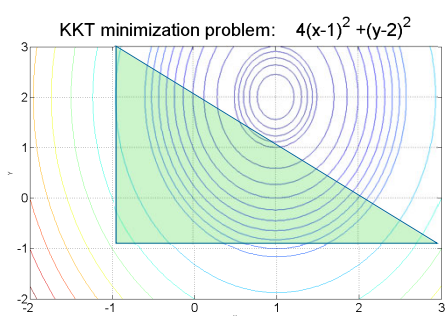
• If solution lies at the exterior, the condition $\nabla f(\mathbf{x}^*) = 0$ does *not* apply. Some constraints are *active*.

**Example**: Minimize

$f(x) = 4(x-1)^2 + (y-2)^2$

with constraints:

$x + y \leq 2$; $x \geq -1$ & $y \geq -1$




– We cannot get any more improvement if for $\mathbf{x}^*$ there does *not* exist a vector **d** that is both a descent direction *and* a feasible direction.

– In other words: the possible feasible directions do not intersect the possible descent directions *at all*.

## 11.4 Mathematical Form

- A vector **d** that is both descending and feasible cannot exist if $-\nabla f = \Sigma\, \mu_i\, (\nabla g_i)$ (with $\mu_i \geq 0$) for all **active** constraints $i \in I$.
  - This can be rewritten as $0 = \nabla f + \Sigma\, \mu_i\, (\nabla g_i)$
  - This condition is correct IF feasibility is defined as $g(\mathbf{x}) \leq 0$.
  - If feasibility is defined as $g(\mathbf{x}) \geq 0$, then this becomes
    $-\nabla f = \Sigma\, \mu_i\, (-\nabla g_i)$
- Again, this only applies for the I ***active*** constraints.
- Usually the **inactive** constraints are included, but the condition $\mu_j\, g_j = 0$ (with $\mu_j \geq 0$) is added for all inactive constraints $j \in J$.
  - This is referred to as the *complimentary slackness* condition.

21

## 11.4 Mathematical Form

• Note that the slackness condition is equivalent to stating that $\mu_j = 0$ for **inactive** constraints -i.e., zero price for non-binding constraints!

• That is, each inequality constraint is either **active**, and in this case it turns into equality constraint of the Lagrange type, or **inactive**, and in this case it is void and does not constrains the solution.

• Note that I + J = M, the total number of (inequality) constraints.

• Analysis of the constraints can help to rule out some combinations. However, in general, a 'brute force' approach in a problem with *J* inequality constraints must be divided into $2^J$ cases. Each case must be solved independently for a minima, and the obtained solution (if any) must be checked to comply with the constrains. A lot of work!

## 11.4 Necessary KKT Conditions

For the problem:

Min $f(\mathbf{x})$

s.t. $g(\mathbf{x}) \le 0$

(n variables, M constraints)

The necessary conditions are:

| | | |
|---|---|---|
| $\nabla f(x) + \Sigma\, \mu_i\, \nabla g_i(x) = 0$ | | (optimality) |
| $g_i(\mathbf{x}) \le 0$ | for i = 1, 2, ..., M | (primary feasibility) |
| $\mu_i\, g_i(\mathbf{x}) = 0$ | for i = 1, 2, ..., M | (complementary slackness) |
| $\mu_i \ge 0$ | for i = 1, 2, ..., M | (non-negativity, dual feasibility) |

Note that the first condition gives n equations.

23

## 11.4 Necessary KKT Conditions - Example

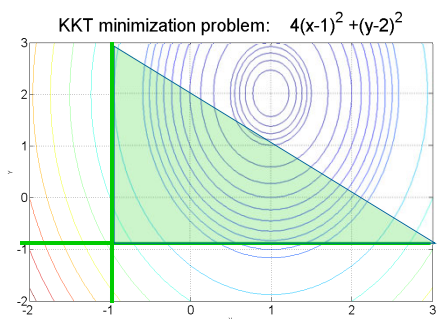
**Example**: Let's minimize

$f(x) = 4(x-1)^2 + (y-2)^2$

with constraints:

$x + y \le 2;\quad x \ge -1\ \&\ \ y \ge -1$

Form the Lagrangian:




$$L(\underline{x},\underline{\lambda},\underline{\mu}) = f(\underline{x}) + \sum_{k=1}^{m} \lambda_k h_k(\underline{x}) + \sum_{j=1}^{\mu} \mu_j g_j(\underline{x})$$

$$L(\underline{x},\underline{\lambda},\underline{\mu}) = 4(x-1)^2 + (y-2)^2 + \mu_1(x+y-2) + \mu_2(x+1) + \mu_3(y+1)$$

There are 3 inequality constraints, each can be chosen active/non-active: 8 possible combinations. But, the 3 constraints together: x+y=2 & x=-1 & y=-1 have no solution, and a combination of any two of them yields a single intersection point.

## 11.4 Necessary KKT Conditions - Example

The general case is:

$$L(\underline{x},\underline{\lambda},\underline{\mu}) = 4(x-1)^2 + (y-2)^2 + \mu_1(x+y-2) + \mu_2(x+1) + \mu_3(y+1)$$

We must consider all the combinations of active / non active constraints:

(1) $x+y=2 \quad \Rightarrow \quad L(x,y,\mu) = 4(x-1)^2 + (y-2)^2 + \mu(x+y-2)$

(2) $x=-1 \quad \Rightarrow \quad L(x,y,\mu) = 4(x-1)^2 + (y-2)^2 + \mu(x+1)$

(3) $y=-1 \quad \Rightarrow \quad L(x,y,\mu) = 4(x-1)^2 + (y-2)^2 + \mu(y+1)$

(4) $x+y=2 \quad \text{and} \quad x=-1 \quad \Rightarrow \quad (x,y)=(-1,3)$

(5) $x+y=2 \quad \text{and} \quad y=-1 \quad \Rightarrow \quad (x,y)=(3,-1)$

(6) $x=-1 \quad \text{and} \quad y=-1 \quad \Rightarrow \quad (x,y)=(-1,-1)$

(7) $x+y=2 \quad \text{and} \quad x=-1 \quad \text{and} \quad x=-1 \Rightarrow \quad (x,y)=\varnothing$

(8) Unconstrained: $\quad L(x,y) = 4(x-1)^2 + (y-2)^2$

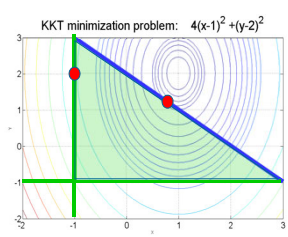
## 11.4 Necessary KKT Conditions - Example

(1) $L(x,y,\mu) = 4(x-1)^2 + (y-2)^2 + \mu(x+y-2)$

$$\left.\begin{aligned} \frac{\partial L}{\partial \mu} &= x+y-2=0 \\ \frac{\partial L}{\partial x} &= 8x-8+\mu=0 \\ \frac{\partial L}{\partial y} &= 2y-4+\mu=0 \end{aligned}\right\} \Rightarrow \left.\begin{aligned} x &= 2-y \\ \mu &= 8-8x \\ 2y-4+8-8(2-y) &= 0 \end{aligned}\right\} \Rightarrow \begin{aligned} (x,y) &= (0.8,1.2) \\ f(x,y) &= 0.8 \\ \mu &= 1.6>0 \end{aligned}$$

(2) $L(x,y,\mu) = 4(x-1)^2 + (y-2)^2 + \mu(x+1)$

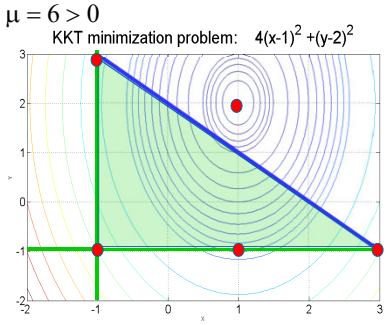
$$\left.\begin{aligned} \frac{\partial L}{\partial \mu} &= x+1=0 \\ \frac{\partial L}{\partial x} &= 8x-8+\mu=0 \\ \frac{\partial L}{\partial y} &= 2y-4=0 \end{aligned}\right\} \Rightarrow \left.\begin{aligned} x &= -1 \\ \mu &= 8-8x=16 \\ y &= 2 \end{aligned}\right\} \Rightarrow \begin{aligned} (x,y) &= (-1,2) \\ f(x,y) &= 16 \\ \mu &= 16>0 \end{aligned}$$



## 11.4 Necessary KKT Conditions - Example

(3) $L(x,y,\mu) = 4(x-1)^2 + (y-2)^2 + \mu(y+1)$

$$\left.\begin{aligned} \frac{\partial L}{\partial \mu} &= y+1=0 \\ \frac{\partial L}{\partial x} &= 8x-8=0 \\ \frac{\partial L}{\partial y} &= 2y-4+\mu=0 \end{aligned}\right\} \Rightarrow \left.\begin{aligned} y &= -1 \\ x &= 1 \\ \mu &= 6 \end{aligned}\right\} \Rightarrow \begin{aligned} &(x,y)=(1,-1) \\ &f(x,y)=9 \\ &\mu = 6 > 0 \end{aligned}$$




(4) $(x,y)=(-1,3);\quad f(x,y)=17$

(5) $(x,y)=(3,-1);\quad f(x,y)=25$

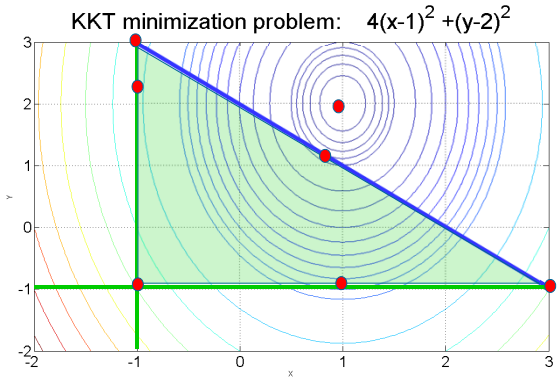
(6) $(x,y)=(-1,-1);\quad f(x,y)=25$

(7) $(x,y)=\varnothing$

(8) $(x,y)=(1,2);\quad f(x,y)=0 \qquad x+y=3>2$ - beyond the range

## 11.4 Necessary KKT Conditions - Example

Finally, we compare among the 8 cases we have studied: case (7) resulted was over-constrained and had no solutions, case (8) violated the constraint $x + y \leq 2$. Among the cased (1)-(6), it was case (1) $(x,y)=(0.8,1.2);\quad f(x,y)=0.8$, yielding the lowest value of $f(x, y)$.



## 11.4 Necessary KKT Conditions: General Case

- For the general case (n variables, M Inequalities, L equalities):

Min $f(\mathbf{x})$ s.t.

$g_i(\mathbf{x}) \le 0$ for $i = 1, 2, ..., M$

$h_j(\mathbf{x}) = 0$ for $J = 1, 2, ..., L$

- In all this, the assumption is that $\nabla g_j(\mathbf{x}^*)$ for j belonging to **active** constraints and $\nabla h_k(\mathbf{x}^*)$ for $k = 1, ..., K$ are linearly independent. This is referred to as *constraint qualification.*

- The necessary conditions are:

$\nabla f(\mathbf{x}) + \Sigma\, \mu_i\, \nabla g_i(\mathbf{x}) + \Sigma\, \lambda_j\, \nabla h_j(\mathbf{x}) = 0$ (optimality)

$g_i(\mathbf{x}) \le 0$ for $i = 1, 2, ..., M$ (primary feasibility)

$h_j(\mathbf{x}) = 0$ for $j = 1, 2, ..., L$ (primary feasibility)

$\mu_i\, g_i(\mathbf{x}) = 0$ for $i = 1, 2, ..., M$ (complementary slackness)

$\mu_i \ge 0$ for $i = 1, 2, ..., M$ (non-negativity, dual feasibility)

(Note: $\lambda_j$ is unrestricted in sign)

## 11.4 Necessary KKT Conditions (if $g(x) \ge 0$)

- If the definition of feasibility changes, the optimality and feasibility conditions change. For example, $g_i(\mathbf{x}) \ge 0$.Then,

Min $f(\mathbf{x})$ s.t.

$g_i(\mathbf{x}) \ge 0$ for $i = 1, 2, ..., M$

$h_j(\mathbf{x}) = 0$ for $J = 1, 2, ..., L$

- The necessary conditions become:

$\nabla f(\mathbf{x}) - \Sigma\, \mu_i\, \nabla g_i(\mathbf{x}) + \Sigma\, \lambda_j\, \nabla h_j(\mathbf{x}) = 0$ (optimality)

$g_i(\mathbf{x}) \ge 0$ for $i = 1, 2, ..., M$ (feasibility)

$h_j(\mathbf{x}) = 0$ for $j = 1, 2, ..., L$ (feasibility)

$\mu_i\, g_i(\mathbf{x}) = 0$ for $i = 1, 2, ..., M$ (complementary slackness)

$\mu_i \ge 0$ for $i = 1, 2, ..., M$ (non-negativity, dual feasibility)

30

## 11.4 Restating the Optimization Problem

- Kuhn Tucker Optimization Problem:

Find vectors $\mathbf{x}_{(Nx1)}$, $\boldsymbol{\mu}_{(1xM)}$ and $\boldsymbol{\lambda}_{(1xK)}$ that satisfy:

| | | |
|---|---|---|
| $\nabla f(\mathbf{x}) + \Sigma\, \mu_i\, \nabla g_i(\mathbf{x}) + \Sigma\, \lambda_j\, \nabla h_j(\mathbf{x}) = 0$ | | (optimality) |
| $g_i(\mathbf{x}) \leq 0$ | for $i = 1, 2, ..., M$ | (feasibility) |
| $h_j(\mathbf{x}) = 0$ | for $j = 1, 2, ..., L$ | (feasibility) |
| $\mu_i\, g_i(\mathbf{x}) = 0$ | for $i = 1, 2, ..., M$ | (complementary slackness condition) |
| $\mu_i \geq 0$ | for $i = 1, 2, ..., M$ | (non-negativity) |

- If $\mathbf{x}^*$ is an optimal solution to NLP, then there exists a $(\boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ such that $(\mathbf{x}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ solves the Kuhn–Tucker problem.
- The above equations not only give the necessary conditions for optimality, but also provide a way of finding the optimal point.

31

## 11.4 KKT Conditions: Limitations

- Necessity theorem helps identify points that are not optimal. A point is not optimal if it does not satisfy the Kuhn–Tucker conditions.
- On the other hand, not all points that satisfy the Kuhn-Tucker conditions are optimal points.
- The Kuhn–Tucker *sufficiency theorem* gives conditions under which a point becomes an optimal solution to a single-objective NLP.

32

## 11.4 KKT Conditions: Sufficiency Condition

- Sufficient conditions that a point x* is a strict local minimum of the classical single objective NLP problem, where f, $g_j$, and $h_k$ are twice differentiable functions are that

  1) The necessary KKT conditions are met.

  2) The Hessian matrix $\nabla^2 L(\mathbf{x}^*) = \nabla^2 f(\mathbf{x}^*) + \Sigma\mu_i \nabla^2 g_i(\mathbf{x}^*) + \Sigma\lambda_j \nabla^2 h_j(\mathbf{x}^*)$ is *positive definite* on a subspace of $R^n$ as defined by the condition:

     $\mathbf{y}^T \nabla^2 L(\mathbf{x}^*)\, \mathbf{y} \geq 0$ is met for every vector $\mathbf{y}_{(1xN)}$ satisfying:

     $\nabla g_j(\mathbf{x}^*)\, \mathbf{y} = 0$ for j belonging to $I_1 = \{ j \mid g_j(\mathbf{x}^*) = 0, u_j^* > 0\}$ (active constraints)

     $\nabla h_k(\mathbf{x}^*)\, \mathbf{y} = 0$ for k = 1, ..., K

     $\mathbf{y} \neq 0$

33

## 11.4 KKT Sufficiency Theorem (Special Case)

- Consider the classical single objective NLP problem.

  minimize $f(\mathbf{x})$ s.t.

  $g_j(\mathbf{x}) \leq 0$ for j = 1, 2, ..., J

  $h_k(\mathbf{x}) = 0$ for k = 1, 2, ..., K

  where $\mathbf{x} = (x_1, x_2, ..., x_N)$

- Features of special case: objective function $f(\mathbf{x})$ is *convex*, all inequality constraints $g_j(\mathbf{x})$ are *convex* functions for j = 1, ..., J, and the equality constraints $h_k(\mathbf{x})$ for k = 1, ..., K are *linear*.
- Then, the necessary KKT conditions are also sufficient.
- Therefore, in this case, if there exists a solution x* that satisfies the KKT necessary conditions, then $\mathbf{x}^*$ is an optimal solution to the NLP problem.
- In fact, it is a *global* optimum.

## 11.4 KKT Conditions: Closing Remarks

- Kuhn-Tucker Conditions are an extension of Lagrangian function and method.
- *They provide powerful means to verify solutions*
- But there are limitations…
  - Sufficiency conditions are difficult to verify.
  - Practical problems do not have required nice properties.
  - For example, you will have a problems if you do not know the explicit constraint equations.
- If you have a multi-objective (lexicographic) formulation, then it is suggested to test each priority level separately.

## 11.4 KKT Conditions: Example

Minimize $C = (x_1 - 4)^2 + (x_2 - 4)^2$ s.t. $2x_1 + 3x_2 \geq 6;$ $12 - 3x_1 - 2x_2 \geq 0;$ $x_1, x_2 \geq 0$

Form Lagrangian : $L = (x_1 - 4)^2 + (x_2 - 4)^2 + \lambda_1(6 - 2x_1 - 3x_2) + \lambda_2(-12 + 3x_1 + 2x_2)$

F.o.c.:

1) $L_{x_1} = 2(x_1 - 4) - 2\lambda_1 + 3\lambda_2 = 0;$

2) $L_{x_2} = 2(x_2 - 4) - 3\lambda_1 + 2\lambda_2 = 0;$

3) $L_{\lambda_1} = 6 - 2x_1 - 3x_2 \leq 0;$

4) $L_{\lambda_2} = -12 + 3x_1 + 2x_2 \leq 0;$

Case 1: Let $\lambda_2 = 0,$ $L_{\lambda_2} < 0$ (2nd Constraint inactive):

From 1) and 2) $\Rightarrow \lambda_1 = x_1 - 4 = 2/3(x_2 - 4);$

From 3) $\Rightarrow x_1 = 3 - 3/2\, x_2;$

$\Rightarrow 3 - 3/2\, x_2 - 4 = 2/3(x_2 - 4)$ $\Rightarrow x_2^* = 5/3 * (6/13)$

$x_2^* = 30/39 = 10/13,$ $x_1^* = 24/13,$ $\lambda_1 = -28/13 < 0$ (Violates KKT conditions)

Case 2: Let $\lambda_1 = 0,$ $L_{\lambda_1} < 0$ (1st Constraint inactive):

From 1) and 2) $\Rightarrow$ $\lambda_2 = -(2/3)(x_1 - 4) = -(x_2 - 4)$ $\Rightarrow x_1 = 3/2(x_2 - 4) + 4$

From 4) $\Rightarrow x_1 = -2/3\, x_2 + 4$

$\Rightarrow -2/3\, x_2 + 4 = 3/2(x_2 - 4) + 4;$ $(-2/3 - 3/2)x_2^* = -6$

36

$x_2^* = 36/13,$ $x_1^* = 84/39 = 28/13,$ $\lambda_2 = 16/13 > 0$ (Meets KKT conditions)

## 11.5 The Constraint Qualification

Example 1 - Irregularities at boundary points

Maximize $\pi = x_1$

subject to $x_2 - (1 - x_1)^3 \le 0$

and $x_1 \,\&\, x_2 \ge 0$

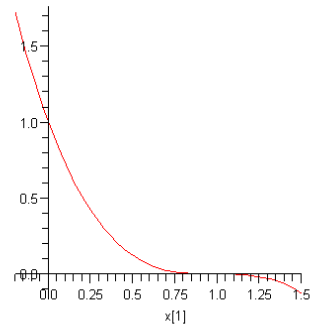
$$L = x_1 + \lambda\left(-x_2 + (1 - x_1)^3\right)$$

$L_{x_1} = 1 - 3\lambda(1 - x_1)^2 \le 0$

$x_1 = 1$ at a max. However,

$L_{x_1} = 1 - 3\lambda(1 - 1)^2 = 1$ when it should equal zero.

Reason : on an inflection point or cusp




37

## 11.5 The Constraint Qualification

Example 2 - Irregularities at the boundary points

Maximize $\pi = x_1$

subject to $x_2 - (1 - x_1)^3 \le 0$

and $2x_1 + x_2 \le 2$

where $x_1 \,\&\, x_2 \ge 0$

$$L = x_1 + \lambda_1\left(-x_2 + (1 - x_1)^3\right) + \lambda_2(2 - 2x_1 - x_2)$$

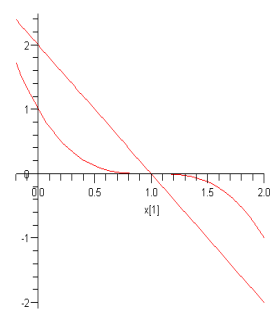
$L_{x_1} = 1 - 3\lambda_1(1 - x_1)^2 - 2\lambda_2 \le 0$

$L_{x_2} = -\lambda_1 - \lambda_2 \le 0$

$L_{\lambda_1} = -x_2 + (1 - x_1)^3 \ge 0$

$L_{\lambda_2} = 2 - 2x_1 - x_2 \ge 0$

$x_1 = 1,\ x_2 = 0,\ \lambda_1 = -\frac{1}{2},\ \lambda_2 = \frac{1}{2}$




38

## 11.5 The Constraint Qualification

Example 3 - The feasible region of the problem contains no cusp

Maximize $\pi = x_2 - x_1^2$

subject to $-\left(10 - x_1^2 - x_2\right)^3 \le 0$ and $-x_1 \ge -2$, where $x_1, x_2 \ge 0$

$$L = x_2 - x_1^2 + \lambda_1\left(10 - x_1^2 - x_2\right)^3 + \lambda_2\left(-2 + x_1\right)$$

$$L_{x_1} = -2x_1 - 6\lambda_1\left(10 - x_1^2 - x_2\right)^2 x_1 + \lambda_2$$

$$L_{x_2} = 1 - 3\lambda_1\left(10 - x_1^2 - x_2\right)^2$$

$$L_{\lambda_1} = \left(10 - x_1^2 - x_2\right)^3$$

$$L_{\lambda_2} = -2 + x_1$$

$$x_1 = 2,\ \ x_2 = 6,\ \ \lambda_1 = \lambda_1, \lambda_2 = 4$$

$L_{x_2} = 1 - 3\lambda_1\left(10 - 2^2 - 6\right)^2 = 1$, when it should equal zero

39