MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Civil and Environmental Engineering

1.731 Water Resource Systems

Lecture 5 & 6, Optimality Conditions, Sept. 21 & 26, 2006

How do we know when a particular **candidate solution** x^* is a local maximum?

Necessary (Kuhn-Tucker) conditions for a candidate solution x^* to be a **local maximum** are:

1. Feasibility
2. Stationarity
3. Inequality Lagrange multipliers
4. Curvature

Preliminaries:
x^* is a **local maximum** if $F(x^*) \geq F(x)$ for all **feasible** x near x^*

m_A^* **active** constraints at x^*: $\quad g_i(x^*) = 0 \quad i \in \mathcal{C}(x^*)$ = **active set**

m_I^* **inactive** constraints at x^*: $\quad g_i(x^*) < 0 \quad i \notin \mathcal{C}(x^*)$

$m_A^* + m_I^* = m$

Form an m_A^* by n matrix with rows the **gradient vectors** $\partial g_i(x^*)/\partial x_j$ of the m_A^* constraint functions active at x^*. If $Rank\,[\partial g_i(x^*)/\partial x_j] = \rho_A^* < m_A^*$ the problem is **degenerate**. Otherwise $\rho_A^* = m_A^*$ and the problem is **non-degenerate**.

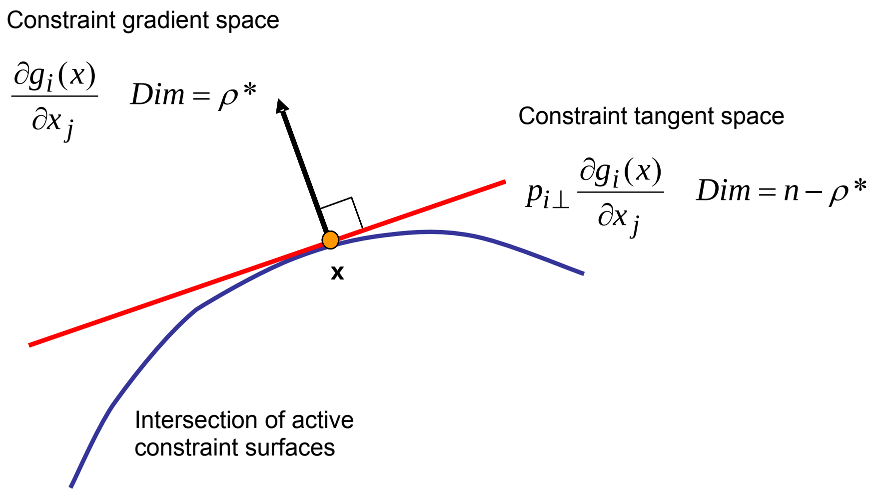
The set of m_A^* constraints active at x^* define an $n - \rho_A^* \geq 0$ dimensional constraint surface in the n dimensional decision space.

Any ρ_A^* **linearly independent** gradient vectors form a basis for a ρ_A^* dimensional **gradient space**. Any $n - \rho_A^*$ **tangent vectors** p_i ($i = 1, \ldots, n - \rho_A^*$) normal to all the gradient vectors form a basis for an $n - \rho_A^*$ dimensional **tangent space**.

Orthogonality condition satisfied by **any** vector p_i in tangent space:

$$p_i \perp \frac{\partial g_j(x^*)}{\partial x_i} \rightarrow p_i \frac{\partial g_j(x^*)}{\partial x_i} = 0 \qquad i \in \mathcal{C}(x^*)$$

The tangent space can be viewed as a plane that intersects the $n - \rho_A^*$ constraint surface at x^*. This plane approximates the constraint surface for x sufficiently close to x^*.



Statement of Necessary Conditions for a Local Maximum:

1. Feasibility

x^* must lie in the feasible region $\mathcal{F}$:

$$g_i(x^*) = 0 \qquad i = 1,\ldots r$$

$$g_i(x^*) \le 0 \qquad i = r+1, \ldots, m$$

2. Stationarity

Objective function gradient at x^* must lie in the constraint gradient space (i.e. it has no projection onto the constraint tangent plane).

For non-degenerate problems this implies:

$$\frac{\partial F(x^*)}{\partial xj} = \lambda_i \frac{\partial g_i(x^*)}{\partial xj} \qquad\qquad i \in \mathcal{C}(x^*)$$

The λ_i are **Lagrange multipliers** for the active constraints at x^*.

If x^* is a **local maximum** this system of n linear equations in the $\rho_A^* = m_A^*$ unknown λ_i's must have a solution (i.e. it must be **consistent**).

For degenerate problems include only $\rho_A^* < m_A^*$ linearly independent constraints and set $\lambda_i = 0$ for the remaining redundant constraints

Adopt convention that $\lambda_i = 0$ for inactive constraints as well as redundant constraints so the stationarity condition can include all constraints:

$$\frac{\partial F(x^*)}{\partial x_j} = \lambda_i \frac{\partial g_i(x^*)}{\partial xj}$$

$$\lambda_i g_i(x^*) = 0 \quad \text{for each } i \in \mathcal{C}(x^*) \text{ (no sum over } i)$$

Define **Lagrangian function** to be:

$$L(x^*,\lambda) = F(x^*) - \lambda_i \; g_i(x^*)$$

Then stationarity condition requires:

$$\frac{\partial L(x^*,\lambda)}{\partial x_j} = 0$$

3. Inequality Lagrange multipliers

If x^* is a **local maximum** then the Lagrange multipliers for all **inequality constraints** active at x^* must be non-negative: $\lambda_i \geq 0, \; i \in \mathcal{C}(x^*)$.

4. Curvature

Projection of Lagrangian onto the constraint tangent space must have a **negative semi-definite Hessian**.

Projection operator is an n by $n - \rho_A^*$ matrix Z_{ik} with columns composed of the $n - \rho_A^*$ constraint tangent space basis vectors. These basis vectors are linearly independent solutions p_i of:

$$p_i \perp \frac{\partial g_i(x^*)}{\partial x_j} \rightarrow p_i \frac{\partial g_j(x^*)}{\partial x_i} = 0$$

$$Z_{ik} = \begin{bmatrix} p_i^1 & \ldots & p_i^{n-\rho_A^*} \end{bmatrix}$$

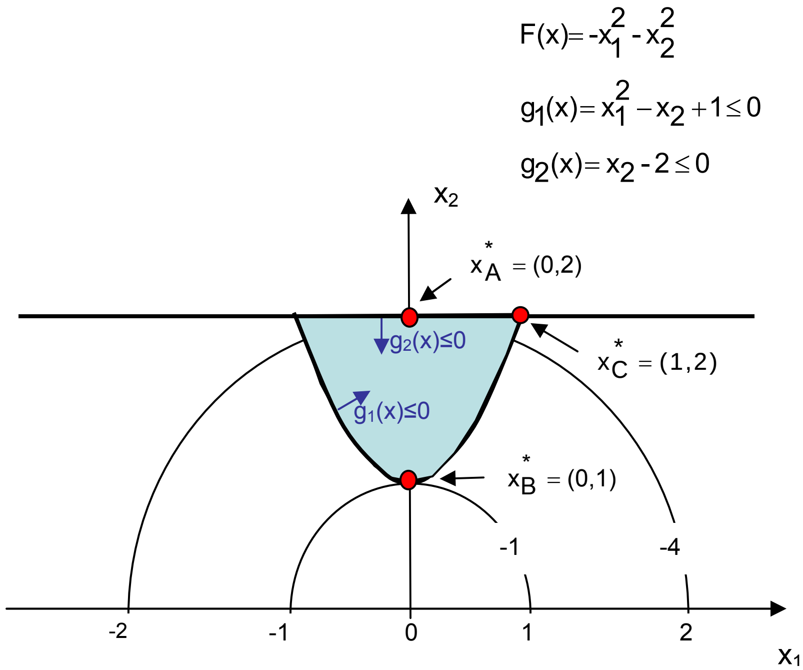
:

Hessian of the projected Lagrangian is W_{kl}:

$$W_{kl} = \frac{\partial^2 L(x^*,\lambda)}{\partial x_j \partial x_i} Z_{ik} Z_{lj} = 0$$

If x^* is a **local maximum**, W_{kl} must be **negative semidefinite** $\rightarrow W_{kl} \leq 0$.

Example:
Consider an two-dimensional example with 2 inequality constraints and 3 candidate solutions
$x_A^* = (0,2)$, $x_B^* = (0,1)$., $x_C^* = (1,2)$



Gradients are:

$$\frac{\partial F}{\partial x_1} = -2x_1 \qquad \frac{\partial F}{\partial x_2} = -2x_2$$

$$\frac{\partial g_1}{\partial x_1} = 2x_1 \qquad \frac{\partial g_1}{\partial x_2} = -1$$

$$\frac{\partial g_2}{\partial x_1} = 0 \qquad \frac{\partial g_2}{\partial x_2} = 1$$

Lagrangian and its Hessian are:

$$L(x,\lambda) = -x_1^2 - x_2^2 - \lambda_1[x_1^2 - x_2 + 1] + \lambda_2[x_2 - 2]$$

$$\frac{\partial L(x,\lambda)}{\partial x_i \partial x_j} = \begin{bmatrix} -2-2\lambda_1 & 0 \\ 0 & -2 \end{bmatrix}$$

Evaluate gradients at candidate solutions

	$\partial F(x^*)/\partial x_j$	$\partial g_1(x^*)/\partial x_j$	$\partial g_2(x^*)/\partial x_j$	m_A^*	ρ_A^*	$n-\rho_A^*$	Z_{ij}
$x_A^* = (0,2)$	$\begin{bmatrix} 0 \\ -4 \end{bmatrix}$	$\begin{bmatrix} 0 \\ -1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	1	1	1	$\begin{bmatrix} a \\ 0 \end{bmatrix}$
$x_B^* = (0,1)$	$\begin{bmatrix} 0 \\ -2 \end{bmatrix}$	$\begin{bmatrix} 0 \\ -1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	1	1	1	$\begin{bmatrix} a \\ 0 \end{bmatrix}$
$x_C^* = (1,2)$	$\begin{bmatrix} -2 \\ -4 \end{bmatrix}$	$\begin{bmatrix} 2 \\ -1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	2	2	0	None

Check necessary conditions for a local maximum at each x^*:

1. Feasibility:
All 3 candidate solutions are feasible.

2. Stationarity
Consider for all active constraints:

	Consistent ?	λ_1	λ_2	W_{kl}
$x_A^* = (0,2)$	Yes	0	-4	$-2a^2 < 0$
$x_B^* = (0,1)$	Yes	+2	0	$-6a^2 < 0$
$x_C^* = (1,2)$	Yes	-1	-5	None

3. Inequality Lagrange Multipliers
Only x_B^* has non-negative Lagrange multipliers for active inequalities.

4. Curvature:
Both x_A^* and x_B^* satisfy the curvature condition. This condition does not apply to x_C^*.

Example Summary:
Only x_B^* is a **local maximum**.since it is the only solution that satisfies all 4 conditions. For this problem x_B^* is also a **global maximum** (why?)

Quick Outline of Derivation:
Derivation of necessary conditions is based on **Taylor series approximations** of $g_i(x)$ and $F(x)$:

An **infinitesimal feasible arc** from x^* to x lies **wholly inside** the feasible region.
Let θ be **distance** along this arc from x^* to x.

$$x^* = x(0) \qquad x = x(\theta)$$

The vector tangent to this arc at $x(0)$ is $\partial x_j(0)/\partial\theta$.

Infinitesimal arcs originating at $x(0)$ are feasible [i.e. $g_i[x(\theta)] = 0$] if the corresponding $\partial x_j(0)/\partial\theta$ lies in the **constraint tangent space**. To see this use a Taylor series expansion of $g_i[x(\theta)]$:

$$g_i[x(\theta)] = g_i[x(0)] + \frac{\partial g_i[x(0)]}{\partial x_j}\frac{\partial x_j(0)}{\partial\theta}\theta + \ldots = 0 \qquad\qquad i \in \mathcal{C}(x^*)$$

The first term on the right is zero because constraint i is active at $x^* = x(0)$.
The second term on the right is zero since $\partial x_j(0)/\partial\theta$ is orthogonal to all the active constraint vectors if it lies in constraint tangent space.

The Taylor series expansion of $F[x(\theta)]$ along an infinitesimal arc is:

$$F[x(\theta)] = F[x(0)] + \frac{\partial F[x(0)]}{\partial \theta}\theta + \frac{\partial^2 F[x(0)]}{\partial \theta^2}\theta^2 + \ldots =$$

$$F[x(\theta)] = F[x^*] + \frac{\partial F[x^*]}{\partial x_i}\frac{\partial x_i(0)}{\partial \theta}\theta + \frac{\partial^2 F[x^*]}{\partial x_i \partial x_j}\frac{\partial x_i(0)}{\partial \theta}\frac{\partial x_j(0)}{\partial \theta}\theta^2 + \ldots$$

If x^* is a local maximum $F(x) = F[x(\theta)]$ must be $\le F[x(0)] = F(x^*)$ for **all** values of θ along the arc. This implies:

1). $$\frac{\partial F[x(0)]}{\partial \theta} = \frac{\partial F[x^*]}{\partial x_j}\frac{\partial x_j(0)}{\partial \theta} = 0$$

2). $$\frac{\partial^2 F[x(0)]}{\partial \theta^2} = \frac{\partial^2 F[x^*]}{\partial x_i \partial x_j}\frac{\partial x_i(0)}{\partial \theta}\frac{\partial x_j(0)}{\partial \theta} \le 0$$

The **stationarity** condition follows from 1). and the **curvature condition** follows from 2), if the requirement that $\partial x_i(0)/\partial \theta$ lies in the constraint tangent space and the definition of the Lagrangian are invoked.

The stationarity condition takes care of feasible arcs that lie in the **tangent space**, which are the only directions that are feasible for **equality** constraints.

If the constraint is an **inequality** the feasible arc may also point **into the feasible region**, away from the tangent space. Directions into the feasible region are defined by::

3). $$\frac{\partial x_j(0)}{\partial \theta}\frac{\partial g_i[x(0)]}{\partial x_j} < 0$$ for i an **inequality** constraint $\in \mathcal{C}(x^*)$

The objective function cannot increase along this feasible arc if x^* is a local maximum. So:

4). $$\frac{\partial F[x(0)]}{\partial x_j}\frac{\partial x_j(0)}{\partial \theta} \le 0$$ for i an **inequality** constraint $\in \mathcal{C}(x^*)$

The **inequality Lagrange multiplier condition** follows from 3) and 4).