Nonlinear Programming

Harold W. Kuhn and Albert W. Tucker

Presented by
E. Alper Yıldırım

School of Operations Research
and Industrial Engineering
Cornell University
Ithaca, NY

February 25, 2000
Harold W. Kuhn (1925-)

• Born in Santa Monica, CA.

• Received PhD in 1950 from Princeton.

• Played a significant role in Nash’s Nobel Prize in 1994.

• Contributions to (non)linear programming, game theory.

• Developed “Hungarian Method” for assignment problem, an early association between combinatorics and continuous optimization.
Albert William Tucker (1906-1995)

• Born in Ontario, Canada.

• Received PhD in 1932 from Princeton.

• Began his career as a topologist.

• Formulated “Prisoner’s Dilemma”.

• John Nash was his PhD student.

• Also visited Cornell.

• Contributions to (non)linear programming and game theory.

• Commitment to teaching of math at colleges.
Some History

- Equality constrained optimization dates back to Euler and Lagrange (18th century).

- William Karush proved the first necessary conditions for general optimization (1939, MS Thesis, University of Chicago, unpublished).

- Fritz John rediscovered one of Karush’s results (1948, rejected by Duke Mathematics Journal, published on 60th birthday anniversary of Courant).

- Kuhn and Tucker rediscovered Karush’s results (1950), with a saddle point motivation.

- Gale declined to participate!
Impact

• Huge!

• Launched the field of nonlinear programming.

• Algorithmic motivation - try to find a point satisfying KKT conditions.

• Sensitivity analysis - notion of shadow prices.

• Generalized duality theory.

• Led to “Convex Analysis” by Rockafellar (1970s).
Nonlinear Programming (NLP)

\[
\begin{align*}
(NLP) & \quad \min_x \ f(x) \\
\text{s.t.} & \quad g(x) \leq 0, \\
& \quad h(x) = 0,
\end{align*}
\]

where \( f : \mathbb{R}^n \rightarrow \mathbb{R} \), \( g : \mathbb{R}^n \rightarrow \mathbb{R}^m \) and \( h : \mathbb{R}^n \rightarrow \mathbb{R}^p \).

NLP is broad!

- \( p = 0 \) inequality constrained
- \( m = 0 \) equality constrained
- \( m = 0, \ p = n, \ f(x) := 0 \) solution of nonlinear equations
- \( p = m = 0 \) unconstrained
- \( h_j(x) = \sin(\Pi x_j) \) integer programming
NLP is far too general to hope to solve.

**Goal:** Characterization of necessary conditions for $x^*$ to be a (local) minimum of $(NLP)$ in the differentiable case.

**Unconstrained Case: First Order Necessary Conditions**

If $x^*$ is a local minimizer of $f$, and $f$ is $C^1$ at $x^*$, then $\nabla f(x^*) = 0$.

Henceforth, assume $f$, $g$ and $h$ are differentiable.
First Order Necessary Conditions

Theorem 1. (W. Karush 1939, F. John 1948) If $x^*$ is a local minimizer of $(NLP)$, then $\exists \tau^* \geq 0, u^* \in \mathbb{R}_+^m, v^* \in \mathbb{R}^p$, not all zero, such that

$$\nabla f(x^*)\tau^* + \nabla g(x^*)u^* + \nabla h(x^*)v^* = 0,$$  \hspace{1cm} (1)

$$g(x^*) \leq 0, \quad h(x^*) = 0,$$  \hspace{1cm} (2)

$$u^* \geq 0, \quad u^*^Tg(x^*) = 0.$$  \hspace{1cm} (3)

Remarks:

- (2) is simply feasibility.

- (3) is complementarity, i.e. $u^*_i g_i(x^*) = 0$.

- Includes the classical Lagrange theorem ($m = 0$) and unconstrained case ($m = p = 0$).
Shortcoming:

\[
\begin{align*}
\min_{x} & \quad -x_1 \\
\text{s.t.} & \quad -(1 - x_1)^3 + x_2 \leq 0, \\
& \quad -x_2 \leq 0.
\end{align*}
\]
**Question:** How does one ensure that \( \tau^* > 0 \)?

**Constraint Qualification (CQ):**

Let \( \mathcal{I} = \{ i \in \{1, \ldots, m \} : g_i(x^*) = 0 \} \), “active set” at \( x^* \).

Let \( D = \{ d : \nabla g_{\mathcal{I}}(x^*)^T d \leq 0, \nabla h(x^*)^T d = 0 \} \).

**Want:** Every \( d \in D \) is a limit of the directions which are tangent to an arc emanating from \( x^* \) which is contained in the feasible region.

If this holds, then we say “**CQ** holds at \( x^* \)”.

**Remarks:**

- There exist sufficient conditions for **CQ** to hold at \( x^* \).
- **CQ** rules out the previous “nasty” case!
First Order Necessary Conditions

**Theorem 2.** (Karush 1939; Kuhn, Tucker 1950) If \( x^* \) is a local minimizer of \((NLP)\) and \(\text{CQ} \) holds at \( x^* \), then \( \exists \ u^* \in \mathbb{R}^m_+ \), \( v^* \in \mathbb{R}^p \) such that

\[
\nabla f(x^*) + \nabla g(x^*) u^* + \nabla h(x^*) v^* = 0, \tag{4}
\]

\[
g(x^*) \leq 0, \quad h(x^*) = 0, \tag{5}
\]

\[
u^* \geq 0, \quad u^*^T g(x^*) = 0. \tag{6}
\]

**Remarks:**

- This is the celebrated “KKT conditions”.
- Note that \( \nabla f(\cdot) \) does not vanish!
Lagrangian Function

Consider the following function (Lagrangian function):

\[ \mathcal{L}(x, u, v) = f(x) + u^T g(x) + v^T h(x) \]

Extend the definition:

\[ \mathcal{L}(x, u, v) = \begin{cases} 
  f(x) + u^T g(x) + v^T h(x) & \text{if } u \geq 0 \\
  -\infty & \text{otherwise} 
\end{cases} \]
Let $S = \{x \in \mathbb{R}^n : g(x) \leq 0, h(x) = 0\}$, the feasible region.

$$\max_{u,v} \mathcal{L}(x, u, v) = \begin{cases} f(x) & \text{if } x \in S \\ +\infty & \text{otherwise} \end{cases}$$

Therefore,

$$(NLP) = \min_{x \in S} f(x) \equiv \min_{x} \max_{u,v} \mathcal{L}(x, u, v)$$
\[ \mathcal{L}(x, u, v) = f(x) + u^T g(x) + v^T h(x) \]

Recall KKT condition (4):

\[ \nabla_x \mathcal{L}(x^*, u^*, v^*) = \nabla f(x^*) + \nabla g(x^*) u^* + \nabla h(x^*) v^* = 0. \]

Therefore, \( x^* \) satisfies the first order necessary conditions to minimize \( \mathcal{L}(\cdot, u^*, v^*) \).

Similarly, recall KKT conditions (5) and (6):

\[ u^* \geq 0, \quad \nabla_u \mathcal{L}(x^*, u^*, v^*) = g(x^*) \leq 0, \]
\[ u^* \nabla_u \mathcal{L}(x^*, u^*, v^*) = 0, \quad \nabla_v \mathcal{L}(x^*, u^*, v^*) = h(x^*) = 0. \]

Therefore, \((u^*, v^*)\) satisfies the first order conditions to maximize \( \mathcal{L}(x^*, \cdot, \cdot) \).
Convex Programming

\((NLP)\) is a convex programming problem if

- \(f\) is convex;
- \(g_i\) are convex, \(i = 1, \ldots, m\);
- \(h_j\) are affine, \(j = 1, \ldots, p\).

Remarks:

- These conditions imply that \(S\) is a convex set.
- The above implication is one-sided.
- Any local minimizer is a global minimizer.
Sufficiency Conditions

**Theorem 3.** *(Kuhn, Tucker 1950)* If *(NLP)* is a convex programming problem, and if \((x^*, u^*, v^*)\) satisfies the KKT conditions, then \(x^*\) is a global minimizer.

**Remarks:**

- Not quite an “if and only if” theorem. *(Need CQ)*

- Strongest first order result (except for linear programming where CQ automatically holds)
On Duality

Recall

$$(NLP) \equiv \min_{x \in S} f(x) = \min_{x} \max_{u,v} \mathcal{L}(x, u, v)$$

Define

$$(NLD) \equiv \max_{u,v} \min_{x} \mathcal{L}(x, u, v)$$

Remarks:

• Might not be useful.

• If $(NLP)$ is convex, then $\mathcal{L}(\cdot, u, v)$ is convex for all $u \geq 0$ and $v$. 
Equivalence

**Theorem 4.** (Kuhn, Tucker 1950) If $(NLP)$ is a convex programming problem and if $(x^*, u^*, v^*)$ satisfies the KKT conditions and CQ holds at $x^*$, then

$$
\mathcal{L}(x^*, u^*, v^*) = \min_x \max_{u,v} \mathcal{L}(x, u, v) = \max_{u,v} \min_x \mathcal{L}(x, u, v).
$$

**Remarks:**

- Follows from Von Neumann’s “Minimax” theorem.

- Reveals the equivalence between $(NLP)$ and the saddle value problem.

- Gives sufficient conditions in order for strong duality to hold.
Quotation

Notable achievements have been recorded in the subjects of convex analysis, the analysis of nonlinear systems, and algorithms to solve optimization problems. This has been possible only because communication has been opened between mathematicians and the potential areas of application, to the benefit of both. . . . the lines of communication between applied fields such as mathematical programming and the practitioners of classical branches of mathematics should be broadened and not narrowed by specialization.

Discussion Questions

1. Why did they not consider the second order conditions?

2. Why was there little contact between different fields?

- Karush was motivated by the calculus of variations with inequality side conditions.
- Kuhn expressed the LP duality as a saddle point property of the Lagrangian.
- Tucker was motivated by the relation between LP and Kirkhoff-Maxwell treatment of electrical networks and duality in topology.
- John’s motivation came from geometrical inequalities.
- All came up with the same result!