ASEN 6020: Optimal Trajectories
Summary of results for Parametric Optimization
D.J. Scheeres, scheeres@colorado.edu

Problem Statement: Given a function $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ defined over a domain $R \subset \mathbb{R}^n$, what are its extreme values (maxima and minima) over the set of $\boldsymbol{x} \in R$ and what $\boldsymbol{x}$ do these occur at.

Existence: If $f(\boldsymbol{x})$ is a continuous function over the domain R and if R is a compact set, then the function will take on a maximum and a minimum value. If either of these conditions are violated, a maximum or a minimum may not exist. If the function is bounded, however, then at least a supremum and an infimum will exist.

Stationary Point: A stationary point is one where the first order variation of the function vanishes for an arbitrary variation of the state, assuming no local constraint. The variation of a function in the interior of R is defined as $\delta f(\boldsymbol{x}) = f_{\boldsymbol{x}} \cdot \delta \boldsymbol{x}$, where $f_{\boldsymbol{x}} = \frac{\partial f}{\partial \boldsymbol{x}}$ and $\delta \boldsymbol{x}$ is the variation of the state.

Corner Point: A corner point is one where the gradient of the cost function $f_{\boldsymbol{x}}$ is undefined along at least one state direction x_i.

Necessary Condition: A necessary condition for a candidate point $\boldsymbol{x}^*$ to yield an extremal value for f is that it be a stationary point, be a corner point, or lie in the boundary of the compact set R.

Global Sufficiency Condition: A sufficient condition for a candidate point $\boldsymbol{x}^*$ to be a global minima are that it satisfy the following inequality: $f(\boldsymbol{x}^* + \delta \boldsymbol{x}) \geq f(\boldsymbol{x}^*), \forall (\boldsymbol{x}^* + \delta \boldsymbol{x}) \in R$. To be a global maxima replace the inequality with $\leq$. To be the unique minima/maxima replace the inequalities with $>$ / $<$, respectively.

Gâteaux Derivative: The Gâteaux derivative (directional derivative) of a function $f(\boldsymbol{x})$ is defined as

$$df_+(\boldsymbol{x}, \boldsymbol{u}) = \lim_{\epsilon \to 0^+} \frac{[f(\boldsymbol{x} + \epsilon \boldsymbol{u}) - f(\boldsymbol{x})]}{\epsilon}$$

where $\boldsymbol{u} \in \mathbb{R}^n$ and $\boldsymbol{x} + \epsilon \boldsymbol{u} \in R$.

Local Sufficiency Conditions: Sufficient conditions for a candidate point $\boldsymbol{x}^*$ to be an extremal are that it satisfy the necessary conditions and satisfy a specific set of local conditions as a function of its properties at the candidate point. There are three variations: it is a smooth interior point, it is a corner interior point, it is a boundary point.

Smooth Interior Point: The sufficient condition for a candidate point to be a minimum (maximum) is that $f_{\boldsymbol{x}\boldsymbol{x}}$ evaluated at the candidate point be positive (negative) definite. If it is indefinite (i.e., has positive and negative eigenvalues) then the point is not an extremum at all but is a saddle point. If it is semi-definite, meaning that there are zero eigenvalues, higher-order partials must be probed along the zero-eigenvector direction.

To show this, the global, unique sufficiency condition can be expanded about the candidate point as

$$\begin{aligned} f(\boldsymbol{x}^* + \delta\boldsymbol{x}) &= f(\boldsymbol{x}^*) + f_{\boldsymbol{x}} \cdot \delta\boldsymbol{x} + \frac{1}{2!}\delta\boldsymbol{x} \cdot f_{\boldsymbol{x}\boldsymbol{x}} \cdot \delta\boldsymbol{x} + \ldots \\ &= f(\boldsymbol{x}^*) + \frac{1}{2!}\delta\boldsymbol{x} \cdot f_{\boldsymbol{x}\boldsymbol{x}} \cdot \delta\boldsymbol{x} + \ldots \\ &> f(\boldsymbol{x}^*) \end{aligned}$$

which is reduced to

$$\frac{1}{2!}\delta\boldsymbol{x} \cdot f_{\boldsymbol{x}\boldsymbol{x}} \cdot \delta\boldsymbol{x} + \ldots \quad > \quad 0$$

for a minimum and < 0 for a maximizer.

Corner Interior Points: The necessary condition is then that the candidate point is a corner point of the function.

The sufficient condition for the candidate point to be a minimum (maximum) is that $df_+(\boldsymbol{x}^*, \boldsymbol{u}) \geq (\leq) 0 \ \forall \boldsymbol{u} \in \mathbb{R}^n$.

Continuity of the function f ensures that the Gâteaux derivative can be defined along any direction $\boldsymbol{u}$. If the derivative is not definite in all directions, then it is simple to prove that the function is "saddle-like" and not extremal at the candidate point.

Boundary Points: It is not required that the first variation of the function at a boundary point be zero for the boundary point to be an extremal.

The sufficient condition for the candidate boundary point to be a minimum (maximum) is that $df_+(\boldsymbol{x}^*, \boldsymbol{u}) \geq (\leq) 0 \ \forall \boldsymbol{u}$ such that $\boldsymbol{x}^* + \epsilon\boldsymbol{u} \in R$.

A different way to phrase this is that $\delta f \geq (\leq) 0 \ \forall \delta\boldsymbol{x}$ such that $\boldsymbol{x}^* + \delta\boldsymbol{x} \in R$. The idea is the same: if all allowable variations in the neighborhood of the boundary point increase (decrease) the function, then the function is a minimum (maximum).

Local and Global Sufficiency Conditions: Local sufficiency conditions are global when $f(\boldsymbol{x})$ is a convex function. A continuous function is convex if

$$\begin{aligned} f(\alpha\boldsymbol{x}_0 + (1-\alpha)\boldsymbol{x}_1) \ \leq \ & \alpha f(\boldsymbol{x}_0) + (1-\alpha) f(\boldsymbol{x}_1) \\ & \forall \boldsymbol{x}_0, \boldsymbol{x}_1 \in R \\ & 0 \leq \alpha \leq 1 \end{aligned}$$

Principle of Constraint: For a given cost function evaluated at an extremal point $\boldsymbol{x}^*$, $f(\boldsymbol{x}^*)$, the addition of any state constraints to the problem will cause a degradation in the overall extremal values. In particular, assume $f(\boldsymbol{x}^*)$ is a minimum point. If constraints are added, shifting the extremal point from $\boldsymbol{x}^*$ to $\boldsymbol{x}^{\#}$, then $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x}^{\#})$. The converse holds for maxima.

Constrained Optimization Problem Statement: The cost function is a continuous function $f(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ defined over a compact domain $R \subset \mathbb{R}^n$. Assume that there are m equality constraints, $g_i(\boldsymbol{x}) = 0$, $i = 1, 2, \ldots, m$, and l inequality constraints, $h_j(\boldsymbol{x}) \geq 0$, $j = 1, 2, \ldots, l$. Assume that $m + l \leq n$.

The problem statement is: What are the extreme values (maxima and minima) of $f(\boldsymbol{x})$ over the set of $\boldsymbol{x} \in R$ subject to the equality and inequality constraints, and what $\boldsymbol{x}$ do these occur at.

This is often called the "Mathematical Programming Problem" (MPP) or sometimes the "Nonlinear Programming Problem" (NPP). If the cost function has the form $f(\boldsymbol{x}) = \boldsymbol{a} \cdot \boldsymbol{x}$, where $\boldsymbol{a} \in \mathbb{R}^n$, then this is the "Linear Programming Problem" (LPP).

Necessary Conditions for Constrained Problems: There are two versions of the necessary conditions, one stronger than the other.

Fritz John Necessary Conditions: Define the Lagrangian

$$L(\boldsymbol{x}, \lambda_0, \boldsymbol{\lambda}, \boldsymbol{\sigma}) = \lambda_0 f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}) + \sum_{j=1}^{l} \sigma_j h_j(\boldsymbol{x})$$

where $\boldsymbol{x} \in R \subset \mathbb{R}^n$, $\lambda_0 \in \mathbb{R}$, $\boldsymbol{\lambda}, \boldsymbol{g} \in \mathbb{R}^m$ and $\boldsymbol{\sigma}, \boldsymbol{h} \in \mathbb{R}^l$.
Then, $\boldsymbol{x}^*$ is an extremal of the MPP if there exists $(\lambda_0, \boldsymbol{\lambda}, \boldsymbol{\sigma}) \neq \boldsymbol{0}$, $\lambda_0 \geq 0$, $\boldsymbol{\sigma} \leq \boldsymbol{0}$ such that

$$\begin{aligned} \left.\frac{\partial L}{\partial \boldsymbol{x}}\right|_* &= \boldsymbol{0} \\ \boldsymbol{g}(\boldsymbol{x}^*) &= \boldsymbol{0} \\ \boldsymbol{h}(\boldsymbol{x}^*) &\geq \boldsymbol{0} \end{aligned}$$

Note, in the Fritz John Conditions the solution can sometimes have $\lambda_0 = 0$. The next set of necessary conditions remove this possibility.

Karush-Kuhn-Tucker (KKT) Necessary Conditions: If the gradients $\partial g_i / \partial \boldsymbol{x}$ and $\partial h_j / \partial \boldsymbol{x}$ are linearly independent for all $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, l$, then $\lambda_0 = 1$ in the Fritz John Necessary Conditions with all the remaining conditions still applying.

Sufficiency Conditions for Constrained Problems: Again we can have local sufficient conditions defined in general or global sufficient conditions defined if additional qualifications are placed on the problem. Sufficiency conditions again essentially reduce to the statement that $\Delta L = L(\boldsymbol{x}^* + \delta\boldsymbol{x}, \boldsymbol{\lambda}^* + \delta\boldsymbol{\lambda}, \boldsymbol{\sigma}^* + \delta\boldsymbol{\sigma}) - L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\sigma}^*) \geq 0$ for all admissible $\delta\boldsymbol{x}$.

Local Sufficiency Conditions: The local sufficiency conditions essentially consider a second order expansion of ΔL about the necessary condition solution to determine if there are any neighboring local solutions that yield a non-definite cost. Assume the KKT Conditions apply to the MPP. Then a solution set $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\sigma}^*)$ to the KKT Conditions is a local extremal if

$$\left.\left[\frac{\partial^2 f}{\partial \boldsymbol{x}^2} + \frac{\partial^2 (\boldsymbol{\lambda} \cdot \boldsymbol{g})}{\partial \boldsymbol{x}^2} + \frac{\partial^2 (\boldsymbol{\sigma} \cdot \boldsymbol{h})}{\partial \boldsymbol{x}^2}\right]\right|_*$$

is a definite matrix $\in \mathbb{R}^{n\times n}$, where $\boldsymbol{\lambda}\cdot\boldsymbol{g} = \sum_{i=1}^{m}\lambda_i g_i$ and $\boldsymbol{\sigma}\cdot\boldsymbol{h} = \sum_{j=1}^{l}\sigma_j h_j$.

Global Sufficiency Conditions: To show that a local extremum is a global extremum, qualifications must be placed on both the cost function and on the constraints.

Assume the KKT Conditions apply to the MPP. Then a solution set $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\sigma}^*)$ to the KKT Conditions is a global extremal if:

- f, $\boldsymbol{g}$ and $\boldsymbol{h}$ are all C^1, meaning that they are continuous and their first derivatives are all well defined.
- f and $\boldsymbol{h}$ are convex functions.
- The equality constraints are affine, meaning $\boldsymbol{g} = \boldsymbol{A}\cdot\boldsymbol{x} + \boldsymbol{b}$ where $\boldsymbol{A} \in \mathbb{R}^{m\times n}$

These conditions ensure that the constraint set is convex and that an extremal is a global extremal.