Unconstrained Problems

First-Order Necessary Conditions

Definition: Relative Minimum

A point $x_0 \in \Omega \subset \mathbb{R}^n$ is a relative minimum point of f over Ω if there is an $\varepsilon > 0$ such that $f(x) \geq f(x_0)$ for all $x \in \Omega$ such that $|x - x_0| < \varepsilon$. If $f(x) > f(x_0)$, x_0 is a strict relative minimum point of f over Ω.

Definition: Global Minimum

A point $x_0 \in \Omega \subset \mathbb{R}^n$ is a global minimum point of f over Ω if there is an $\varepsilon > 0$ such that $f(x) \geq f(x_0)$ for all $x \in \Omega$. If $f(x) > f(x_0)$, x_0 is a strict global minimum point of f over Ω.

Proposition: First-Order Necessary Conditions

Let $\Omega \subset \mathbb{R}^n$ and let $f : \Omega \to \mathbb{R} \in C^1$. If x_0 is a relative minimum point of f, then for any feasible direction d, . $\nabla f(x_0) d \geq 0$

Corollary

Let $\Omega \subset \mathbb{R}^n$ and let $f : \Omega \to \mathbb{R} \in C^1$. If x_0 is a relative minimum point of f and if x_0 is an interior point of Ω, then $\nabla f(x_0) = 0$.

Second-Order Conditions

Proposition: Second-Order Necessary Conditions

Let $\Omega \subset \mathbb{R}^n$ and let $f : \Omega \to \mathbb{R} \in C^2$. If x_0 is a relative minimum point of f, then for any feasible direction d,

1. $\nabla f(x_0) d \geq 0$,
2. if $\nabla f(x_0) d = 0$ then $d^T \nabla^2 f(x_0) d \geq 0$.

Proposition: Second-Order Necessary Conditions

Let $\Omega \subset \mathbb{R}^n$ and let $f : \Omega \to \mathbb{R} \in C^2$. If x_0 is a relative minimum point of f and if x_0 is an interior point of Ω, then

1. $\nabla f(x_0) = 0$,
2. $d^T \nabla^2 f(x_0) d \geq 0$ for all $d \in \mathbb{R}^n$.

Proposition: Second-Order Sufficient Conditions

Let $\Omega \subset \mathbb{R}^n$ and let $f : \Omega \to \mathbb{R} \in C^2$. Let x_0 is be an interior point of Ω. Suppose in addition that

1. $\nabla f(x_0) = 0$,
2. $F(x_0) \stackrel{\text{def}}{=} \nabla^2 f(x_0)$ is positive definite (i.e. $d^T \nabla^2 f(x_0) d \geq 0 \ \ \forall d \in \mathbb{R}^n$).

Then x_0 is a strict relative minimum point of f.

Convexity

Definition: Convex Set

A set S is convex if for all $x, y \in S$ and $0 \leq \alpha \leq 1$, $\alpha x + (1 - \alpha) y \in S$.
Note: Geometrically, if $x, y \in S$, then the closed line segment $y + \alpha(x - y)$ joining x and y is in S.

Example
If A and B are convex sets, then $A \cap B$ is a convex set.

Definition: Convex Hull
Consider all convex sets A containing a set S (not necessarily convex). The convex hull is $H(S) = \bigcap_{A \supseteq S} A$.

Definition: Convex Function
Let $\mathrm{Dom}(f) \subset \mathbb{R}^n$ be a convex set and let $f : \mathrm{Dom}(f) \to \mathbb{R}$. Then f is convex if for $0 \leq \alpha \leq 1$,
$f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$.

Proposition
Let $f \in C^1$. Assume Ω is a convex set. Then f is convex on Ω if and only if $f(y) \geq f(x) + \nabla f(x)(y-x)$ for all $x, y \in \Omega$.

Proposition
Let $f \in C^2$. Assume Ω is a convex set containing an interior point. Then f is convex on Ω if and only if Hess f is positive semidefinite on the interior of Ω.

Claim
If f is convex on a convex set Ω, then $S = \{x \mid f(x) \leq c, c \text{ a constant}\}$ is also convex.

Definition: Concave
f is concave if $-f$ is convex.

Minimization and Maximization of Convex Functions

Theorem
Let f be convex on a convex set Ω. Then any relative minimum of f is global.

Definition: Extreme Point
A point $x \in \Omega$ is an extreme point if there are no points $x_1, x_2 \in \Omega$ and $0 \leq \alpha \leq 1$ such that $x = \alpha x_1 + (1-\alpha) x_2$.

Theorem
Let f be convex on a closed and bounded (compact) convex set Ω. If f has a maximum, this much be achieved at an extreme point of Ω.

Numerical Methods
Iterative methods that produce a sequence of points converging to the mininum.

Method of Steepest Decent
- Require a starting approximation x_0.
- Direction of steepest decent is $-\nabla f(x_0)$. Want to find t such that $f(x_0 - t \nabla f(x_0))$ is minimized.
- It suffices to consider the vertical plane through x_0 containing the direction vector $\nabla f(x_0)$, i.e. it suffices to study a

real valued function of one variable.

- Choose x_1 so $f(x_1)$ is minimized. This can be done with line search methods.
- Once minimum is found, obtain another direction of steepest decent and repeat.
- If it fails, $\{x_k\}$ is nowhere near correct values.

Newton's Line Search Formula

$x_{k+1}=x_k-\dfrac{f'(x_k)}{f''(x_k)}$, i.e. Newton's method for computing $f'(x)=0$.

Speed of Convergence

Definition: Order of Convergence

Let $r_k\rightarrow r$. Consider the p's such that $\limsup_{k\rightarrow\infty}\dfrac{|r_{k+1}-r|}{|r_k-r|^p}$ is finite. Then $\operatorname{lub}\{p\}$ is the order of convergence.

Constrained Minimization

Constraints

Minimized f on Ω defined by $\begin{matrix} h_1(x)=0 \\ \vdots \\ h_m(x)=0 \end{matrix}$ or by $\begin{matrix} g_1(x)\le 0 \\ \vdots \\ g_p(x)\le 0 \end{matrix}$; we require these to be "independent".

Definition: Regular Point

A point $a\in\mathbb{R}^n$ is regular if $\nabla h(a)=\begin{bmatrix} \nabla h_1(a) \\ \vdots \\ \nabla h_m(a) \end{bmatrix}$ have linearly independent rows.

First-Order Necessary Conditions (Equality Constraints)

Lemma

Let a be a regular point of the constraints $h(x)=0$ and a local extremum point (minimum or maximum) of f subject to these constraints. Then for all $y\in\mathbb{R}^n$ satisfying $\nabla h(a)y=0$ must also satisfy $\nabla f(a)y=0$.

Theorem

Let a be a regular point of the constraints $h(x)=0$ and a local extremum point (minimum or maximum) of f subject to these constraints. Then there exists $\lambda\in\mathbb{R}^n$ such that $\nabla f(a)+\lambda^T\nabla h(a)=0$.

Second Order Conditions

Theorem: Second Order Necessary Conditions

Suppose a is a local minimum for f subject to $h(x)=0$ and that a is a regular point of these conditions. Then there exists $\lambda \in \mathbb{R}^m$ such that $\nabla f(a)+\lambda^T \nabla h(a)=0$ (first-order conditions satisfied). If we donote the tangent space $M=\{y | \nabla h(a) y=0\}$, then the matrix $L(a)=F(a)+\lambda^T H(a)$ is positive semidefinite on M, i.e. $y^T L(a) y \geq 0, \ \forall y \in M$.
Note: $\lambda^T H(a)=\text{Hess}(\lambda^T h(a))$.
Note: $L(a)=\text{Hess}(f+\lambda^T h)(a)$. This takes care of the curvature of the constraint space.

Theorem: Second Order Sufficiency Conditions

Suppose there is a point a satisfying $h(a)=0$, and a $\lambda \in \mathbb{R}^m$ such that $\nabla f(a)+\lambda^T \nabla h(a)=0$. Suppose also that the matrix $L(a)=F(a)+\lambda^T H(a)$ is positive definite on $M=\{y | \nabla h(a) y=0\}$. Then a is a strict local minimum of f subject to $h(x)=0$.

Eigenvalues in Tanget Subspaces

Suppose $M=\{y | \nabla h(a) y=0\}$.

- Find a basis in M – find enough linearly independent vectors.
- Use Gram-Schmidt to produce an orthonormal basis $\{e_{1,} \ldots, e_k\}$; we know $k+m=n \Leftrightarrow m=n-k$.
- Construct a projection operator P onto M. If $\langle\,,\rangle$ is the inner product, $Px=\langle x, e_1\rangle e_1+\cdots+\langle x, e_k\rangle e_k$. Turn P into a matrix.
- Consider $P^T L P$ a matrix operating in M. For any $x \in \mathbb{R}^n$, $y=Px \in M$.

Bordered Hessians

$\det \begin{bmatrix} 0 & \nabla h(a) \\ -(\nabla h(a))^T & L-\lambda I \end{bmatrix}=p(\lambda)=0$ is the characteristic polynomial of L. Its roots are the eigenvalues of L restricted to M.

Sensitivity

Theorem: Sensitivity Theorem

Let $f, h \in C^2$. Consider the family of problems with parameter c: minimize $f(x)$ subject to $h(x)=c$. Suppose for $c=0$ there is a local solution a that is a regular point and that, together with its associated Lagrange multiplier vector λ, satisfies the second-order sufficiency conditions for a strict local minimum. Then for every $c \in \mathbb{R}^m$ in a region containing $0 \in \mathbb{R}^m$ there is an $x: \mathbb{R}^m \to \mathbb{R}^m$ depending continuously on c such that $x(0)=a$ and $x(c)$ is a local minimum of the family of problems. Furthermore, $\nabla_c f(x(c))|_{c=0}=-\lambda^T$.

Khun-Tucker Conditions (Inequality Constraints)

Theorem

Constraints: $h(x)=\begin{pmatrix} h_1(x) \\ \vdots \\ h_m(x) \end{pmatrix}=0$ and $g(x)=\begin{pmatrix} g_1(x) \\ \vdots \\ g_p(x) \end{pmatrix} \leq 0$. Suppose a is a local minimum for $f(x)$ and is a regular point with respect to active constraints (ie. h_i's and g_j's such that $g_j(x)=0$). Then there exists $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p, \mu \geq 0$

such that $\begin{cases} \nabla f(a)+\lambda^T \nabla h(a)+\mu^T \nabla g(a)=0 \\ h(a)=0 \\ \mu^T g(a)=0 \end{cases}$.

Control Theory

Introduction

$\max \psi(x(t_f))$ subject to $x(t_0)=x_0$ and $\dot{x}(t)=f(t, x(t), u(t))$.

- t is a variable with $t_0 \le t \le t_f$
- $\dot{x}(t)=\frac{dx}{dt}$.
- $u(t)$ is the controller function. $u(t)\in U$ and u piecewise continuous.
- Both x and u can be vector valued, i.e. $x\in\mathbb{R}^n$, $y\in\mathbb{R}^m$.
- Problem: Choose u so that $\psi(x(t_f))$ is maximized.

Solution Procedure

- Define a new function $\lambda(t)$ where $t_0\le t\le t_f$ and $\lambda\in\mathbb{R}^n$.
- For each t, consider the Hamiltonian $H(t, x, u, \lambda)=\lambda(t)^T f(t, x(t), u(t))$.
- Fix λ, t, x. Find a $u^*(t, x, \lambda)$ that yields $\max_{u\in U} H(t, x, u, \lambda)$. Hope $u^*(t, x, \lambda)$ is piecewise continuous. $u^*(t, x, \lambda)$ produces a solution $x^*(t)$.
- Solve the adjoint differential equation $\dot{\lambda}(t)=-\nabla_x H(t, x, u^*(t, x, \lambda), \lambda)$ with $\lambda(t_f)=\nabla_x \psi(x^*(t_f))$ ($u^*(t, x, \lambda)$ produces a solution $x^*(t)$).

Theorem

$\max \psi(x(t_f))$ subject to $x'=f(f, x, u)$ on $[t_0, t_f]$ with initial condition $g^0(x(t_0))=b^0$ (i.e. initial $x(t_0)$ must like on given starting manifold) and final condition $g^f(x(t_f))=b^f$ (i.e. final $x(t_f)$ must like on given terminal manifold). The controller constraint is $u\in U$ and piecewise-continuous.

Notation:

- Let $T^0 \stackrel{\text{def}}{=} \{x : g^0(x)=b^0\}$ be the starting manifold. Let $x^*\in T^0$ and define $T^0_{an} \stackrel{\text{def}}{=} \{x : \nabla g^0(x^*)x=0\}$ to be the tangent plane.
- Let $T^f \stackrel{\text{def}}{=} \{x : g^f(x)=b^f\}$ be the terminal manifold. Let $x^*\in T^f$ and define $T^f_{an} \stackrel{\text{def}}{=} \{x : \nabla g^f(x^*)x=0\}$ to be the tangent plane.

Define the Hamiltonian $H(t, x, u, p)=p^T f(t, x, u)$. Let $\max_u H(t, x, u, p)=H(t, x, u^*(t, x, p), p)=M(t, x, p)$. Define the adjoint ODE $-p^{T\,\prime}=\nabla_x H(t, x, u^*(t, x, p), p)$.

We look for a solution $\begin{pmatrix} x(t) \\ p(t) \end{pmatrix}$ which satisfies $\begin{cases} p(t_0)\perp T^0_{an}(x(t_0)) \\ [p(t_f)-p_0 * \nabla_x \psi(x(t_f))]\perp T^f_{an}(x(t_f)) \end{cases}$.

For practical purposes, this is equivalent to satisfying $\begin{cases} p(t_0)=\nu^0 \nabla_x g^0(x(t_0)) \\ [p(t_f)-p_0 * \nabla_x \psi(x(t_f))]=\nu^f \nabla_x g^f(x(t_f)) \end{cases}$ where ν^0 and ν^f are Lagrange multipliers.

Attainable Set

Definition: Attainable Set
$x'=f(t,x,u)$. Consider the set of landing points of the solution curves as u varies. This is calle the attainable set at t_f , denoted by $K(t_f)$.

Extremal Solutions
We know that there are solutions which follows $\partial K(t)$. They are called extremal solutions.
To find an extremal solution:

- $x(t_0)=a$ given.
- Choose $p(t_0)$.
- Let $H(x,t,u,p)=p\,f(t,x,u)$. Find $\max\limits_u H$ and obtain $u^*=u^*(t,x,p)$.
- Solve $-p'=\nabla_x H(t,x,u^*,p)$ together with original system $x'=f(t,x,u)$ and initial value $p(t_0)$.

Note: This is an IVP, so solution is unique.
Note: $x(t)$ is an extremal solution, and $p(t)\perp\partial K(t)$ at point $x(t)$.

Dynamic Programming

- Most useful for stochastic control problem.
- Example: Find the path with shortest travel time.

Definition: Value Function
Define a function on the graph with gives the shortest travel time from the node to finish and tells the optimal exit direction. This is called the value function.

Theorem
$\max\left[\Phi(x(t_f))+\int_0^{t_f} f_0(t,x(t),u(t))dt\right]$ subject to $x'(t)=f(t,x(t),u(t))$ with initial condition $x(0)=x_0$ and controller constraint $u:[0,t_f]\to\Omega$ piecewise continuous.
Note that this is in Bolza form. To solve it, notice it is equivalent to the following Mayer problem; Define $y(t)=\int_0^{t_f} f_0(t,x(t),u(t))dt$; then $\begin{pmatrix}x\\y\end{pmatrix}'=\begin{pmatrix}f(t,x,u)\\f_0(t,x,u)\end{pmatrix}$ and solve $\max\left[\psi\begin{pmatrix}x(t_f)\\y(t_f)\end{pmatrix}\right]=\max\left[\Phi(x(t_f))+y(t_f)\right]$. Note that the Hamiltonian is, without loss of generality, $H=f_0(t,x,u)+\lambda f(t,x,u)$.

Theorem: Hamilton-Jacobi-Bellman Equation
Define value function $V(t_0,x_0)=\max\left[\Phi(x(t_f))+\int_0^{t_f} f_0(t,x(t),u(t))\,dt\right]$, with $x'(t)=f(t,x(t),u(t))$ and initial condition $x(0)=x_0$. Then define $H(t,x,u,p)=f_0+p\,f(t,x,u)$. $\max\limits_u H$ to obtain $H^{\max}(t,x,p)$. Then the value function satisfy $\partial_t V+H^{\max}(t,x,\nabla_x V)=0$. When $t=t_f$, $V(t_f,x_f)=\Phi(x_f)$.