CMU 16-745 最优控制笔记

吕翔宇 (Xiangyu Lyu)
lvxiangyu11@gmail.com
TU Darmstadt, Germany
GitHub: lvxiangyu11

2025 年 7 月 2 日

目录

1 引言

这里是 CMU 16-745 最优控制课程的笔记。我将涵盖课程中的重要概念、定理和例子。课程目标

1. 分析动力系统的稳定性。
2. 设计稳定平衡和轨迹的LQR控制器。
3. 使用离线轨迹优化设计非线性系统的轨迹。
4. 使用在线凸优化实现模型预测控制。
5. 了解随机性和模型不稳定性的影响。
6. 无最优模型时直接优化反馈策略。

写作风格备注：（TODO：终稿注释掉!）额外的公式推导使用浅绿色背景框额外的例子使用浅蓝色背景框代码使用浅灰色背景

2 机器人中的数值优化引言

从这一节开始，我们将转向机器人中的数值优化。我们将介绍一些基本的概念和方法，继续前面所论述的最优控制内容。

本课程将围绕以下几个主题展开：

- **数值优化基础**：介绍优化问题的基本形式、约束类型、凸性等核心概念。
- **无约束优化方法**：包括梯度下降法、牛顿法、拟牛顿法等常用算法及其收敛性分析。
- **约束优化方法**：涵盖拉格朗日乘子法、KKT条件、罚函数法、投影法等。
- **凸优化与非凸优化**：讲解凸优化的理论基础及其在机器人中的应用，讨论非凸问题的挑战与常用求解策略。
- **数值优化在机器人中的应用**：如轨迹规划、参数估计、点云配准、路径平滑等实际案例。
- **现代优化工具与软件**：介绍常用的数值优化库和求解器（如CVX、Gurobi、OSQP等），并结合实际问题进行演示。

推荐书目：

- **《最优化：建模算法与理论》**：适合初学者，涵盖线性和非线性优化问题的求解方法。
- **Numerical Optimization**，Nocedal 和 Wright：深入探讨数值优化的理论和算法，包括梯度下降、牛顿法等经典方法。
- **Lectures on Convex Optimization**，Bertsekas：理论清晰，涵盖光滑与非光滑优化。
- **Lectures On Modern Convex Optimization**：分析、算法与工程应用。

2.1 优化问题的基本形式

优化问题通常可以表述为：

$$
\begin{aligned}
\min_{x} \quad & f(x) \\
\text{s.t.} \quad & g(x) \leq 0 \\
& h(x) = 0
\end{aligned}
$$

其中，

- $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$：优化变量
- $f : \mathbb{R}^n \to \mathbb{R}$：目标函数
- $g : \mathbb{R}^n \to \mathbb{R}^{m_g}$：不等式约束函数
- $h : \mathbb{R}^n \to \mathbb{R}^{m_h}$：等式约束函数

最优解 x^* 是在所有满足约束条件的向量中，使目标函数 f 取得最小值的解。其中 $f(x)$ 要满足 1. lower bounded，2. 有 bounded level set, 即 $f(x) \leq \beta$ 的解集是有界的。

在机器人学中，最优化常用于：平滑和映射（用非线形最小二乘），轨迹规划（使用非线形规划），点云配准（半正定规划），时间优化的参数化（二阶Conic Program）。

下面是课程安排：

1. **第一次课：数值优化基础**
   (a) 数学规划与机器人
   (b) 非凸优化中的凸性
   (c) 凸集与凸函数
   (d) 无约束非凸优化
   (e) 线搜索最速梯度下降
   (f) 修正阻尼牛顿法
2. **第二次课：无约束优化进阶**
   (a) 拟牛顿法工程上有效的无约束优化方法
   (b) BFGS 更新
   (c) 强/弱曲率条件
   (d) Li-Fukushima 全局收敛性：介绍Li-Fukushima方法在无约束优化中的全局收敛性理论，分析其与传统牛顿法的异同及在实际问题中的应用。
   (e) 有限记忆 BFGS 轻量化
   (f) Lewis-Overton 非光滑线搜索
   (g) 线性共轭梯度下降绕过大矩阵
   (h) 牛顿-CG 方法
   (i) 应用：平滑导航路径生成
3. **第三次课：约束优化方法**

(a) 约束优化问题的分类与复杂性
(b) 低维线性规划：Seidel 算法
(c) 低维严格凸二次规划
(d) 顺序无约束极小化技术（SUMT）
(e) SUMT：罚函数法与障碍法
(f) SUMT：拉格朗日松弛与Uzawa方法
(g) Karush-Kuhn-Tucker (KKT) 条件
(h) 增广拉格朗日方法（PHR）
(i) 应用1：控制分配问题
(j) 应用2：碰撞距离计算
(k) 应用3：非线性模型预测控制

4. **第四次课：对称锥优化**
   (a) 锥与对称锥
   (b) 对称锥与欧几里得Jordan代数
   (c) 锥的可表性
   (d) 对称锥的增广拉格朗日法
   (e) 半光滑牛顿法
   (f) 应用：时间最优路径参数化
5. **第五次课：问题建模与求解技巧**
   (a) 平滑化技术
   (b) 自由度与伴随法
   (c) 线性求解器的分类与特性
   (d) 项目：密集障碍环境下的安全导航

3

机器人中的数值优化第一章————数值优化基础

3.1 凸集与非凸集

定义 3.1 (凸集). 如果对于任意两个点 $x, y \in C$，以及任意 $\lambda \in [0,1]$，都有 $\lambda x + (1-\lambda)y \in C$，则称 C 是一个凸集。

更一般的，凸集中的所有点的任意凸组合也在该凸集内：

$$\sum \theta_i x_i, \sum \theta_i = 1, \theta_i \leq 0 \tag{1}$$

convex hull：选一个最小的凸集能将原来非凸集包起来。

常见集合的凸性举例

- **超平面（Hyperplane）**：$\{x \mid a^T x = b\}$ 是凸集。因为任意两点的凸组合仍满足线性等式。
- **半空间（Half-space）**：$\{x \mid a^T x \geq b\}$ 是凸集。因为线性不等式的解集是凸的。
- **球体（Sphere）**：$\{x \mid \|x - x_0\| = b\}$ 不是凸集。因为两个在球面上的点的连线一般不在球面上。
- **球（Ball）**：$\{x \mid \|x - x_0\| \leq b\}$ 是凸集。因为范数是凸函数，其下水平集是凸的。
- **多项式零点集（Polynomials）**：$\{f \mid f = \sum_i a_i x^i\}$ 一般不是凸集，除非对系数或函数有特殊限制。

定义 3.2 (锥（Cone）). 如果集合 C 满足：对于任意 $x \in C$ 和任意 $a \geq 0$，都有 $ax \in C$，则称 C 是一个锥集（cone）。

例子 3.3 (二阶锥). 二阶锥（Second-order cone）定义为

$$C_2 = \{(x,t) \mid \|x\| \leq t\} \subset \mathbb{R}^{n+1}$$

其中 $x \in \mathbb{R}^n, t \in \mathbb{R}$。

定义 3.4 (半正定锥（Semi-definite cone）). 半正定锥 S_+^n 定义为所有**对称**且**半正定**的 $n \times n$ 实矩阵的集合：

$$S_+^n = \left\{A \in \mathbb{R}^{n \times n} \mid A = A^T,\, A \succeq 0\right\}$$

[1] 其中 $A \succeq 0$ 表示 A 是半正定的，即对任意 $x \in \mathbb{R}^n$，有 $x^T A x \geq 0$。

注释 3.5. 半正定锥是一个凸锥：若 $A, B \in S_+^n$，且 $a, b \geq 0$，则 $aA + bB \in S_+^n$。

证明. 由于 A, B 都是对称半正定矩阵，$a, b \geq 0$，则 $aA + bB$ 也是对称矩阵。对任意 $x \in \mathbb{R}^n$，有

$$x^T(aA + bB)x = a\, x^T A x + b\, x^T B x \geq 0$$

因为 $x^T A x \geq 0$，$x^T B x \geq 0$，且 $a, b \geq 0$。因此 $aA+bB$ 也是半正定矩阵，即 $aA+bB \in S_+^n$。 □

[1]这里矩阵半正定等定义和判定见矩阵分析课程，此处略。

3.1.1 凸集的基本性质

- **任意个凸集的交集仍为凸集**。即 $\bigcap_i C_i$ 是凸集，只要每个 C_i 都是凸集。
- **凸集的并集一般不是凸集**。只有在所有集合两两包含时，才有可能为凸集。
- **仿射变换保持凸性**。若 C 是凸集，A 为线性变换，b 为向量，则 $AC+b = \{Ax+b \mid x \in C\}$ 也是凸集。
- **凸集的闭包、内部、相对内部仍为凸集**。
- **凸锥的非负缩放仍为凸锥**。
- **凸集的Minkowski和（加法）一般是凸集**。设 C_1, C_2 是凸集，则 $C_1 + C_2 = \{x + y \mid x \in C_1, y \in C_2\}$ 也是凸集。
- **笛卡尔积保持凸性**。若 $A \subset \mathbb{R}^n$、$B \subset \mathbb{R}^m$ 都是凸集，则 $A \times B$ 也是凸集。

  证明. 设 $(a_1, b_1), (a_2, b_2) \in A \times B$，任取 $\lambda \in [0,1]$，则

  $$\lambda(a_1, b_1) + (1-\lambda)(a_2, b_2) = (\lambda a_1 + (1-\lambda)a_2,\ \lambda b_1 + (1-\lambda)b_2)$$

  由于 A 是凸集，$\lambda a_1 + (1-\lambda)a_2 \in A$；$B$ 是凸集，$\lambda b_1 + (1-\lambda)b_2 \in B$。因此

  $$(\lambda a_1 + (1-\lambda)a_2,\ \lambda b_1 + (1-\lambda)b_2) \in A \times B$$

  所以 $A \times B$ 是凸集。 □

- **凸集的Hadamard乘积（逐元素乘法）一般不是凸集**。只有在特殊情况下（如所有元素非负等）才可能保持凸性。

3.1.2 半正定锥的凸性与几何形状

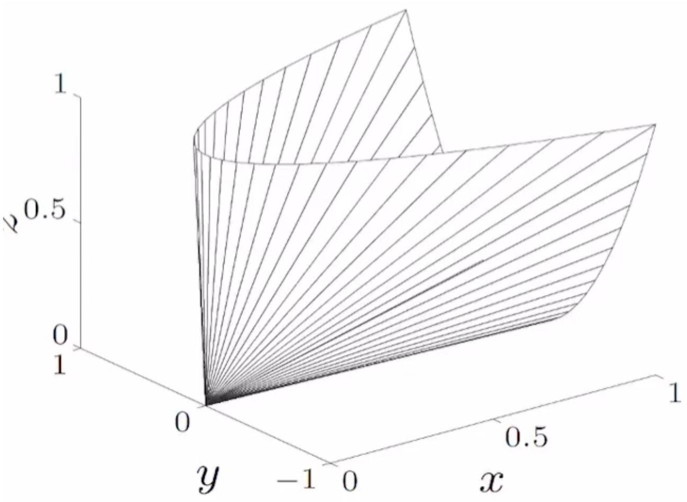

图 1: 半正定锥 S_+^n 的几何形状示意图。

半正定锥的定义与凸性　半正定锥 S_+^n 可以用如下方式刻画：

$$S_+^n = \bigcap_{x \in \mathbb{R}^n} \left\{ A \mid x^T A x \geq 0 \right\}$$

即所有对称矩阵 A，使得对任意 x 都有 $x^T A x \geq 0$。

凸性分析　半正定锥是凸集。因为对于任意 $A, B \in S_+^n$，以及任意 $\theta \in [0,1]$，有

$$x^T \left(\theta A + (1-\theta) B \right) x = \theta x^T A x + (1-\theta) x^T B x \geq 0$$

因此 $\theta A + (1-\theta) B \in S_+^n$。

几何直观　如上图所示，半正定锥在低维空间中呈现“锥形”结构，是一个凸锥。该集合在优化中非常重要，广泛用于半正定规划（SDP）等问题。

3.2　高阶函数

梯度、Jacobian与Hessian

- **梯度（Gradient）**：对于标量函数 $f(x) : \mathbb{R}^n \to \mathbb{R}$，梯度是所有一阶偏导数组成的向量：

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

- **Jacobian矩阵（Jacobian）**：对于向量值函数 $f(x) : \mathbb{R}^n \to \mathbb{R}^m$，Jacobian是所有一阶偏导数组成的 $m \times n$ 矩阵：

$$J_f(x) = \frac{\partial f}{\partial x^T} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

- **Hessian矩阵（Hessian）**：对于标量函数 $f(x) : \mathbb{R}^n \to \mathbb{R}$，Hessian是所有二阶偏导数组成的对称矩阵：

$$H_f(x) = \nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

高阶函数的泰勒展开　设 $f : \mathbb{R}^n \to \mathbb{R}$ 是光滑函数，$x \in \mathbb{R}^n$，其泰勒展开为：

$$f(x) = f(0) + x^T \nabla f(0) + \frac{1}{2} x^T \nabla^2 f(0) x + O(\|x\|^3)$$

该展开在数值优化中用于函数的线性和二次近似。

3.3 高阶函数的信息与微分记号

矩阵与向量的导数类型 对于不同类型的输入输出，常见的导数记号如下表所示：

X	Y	G	记号	名称
$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$f'(x)$	导数
$\mathbb{R}^n$	$\mathbb{R}$	$\mathbb{R}^n$	$\frac{\partial f}{\partial x_i}$	梯度
$\mathbb{R}^n$	$\mathbb{R}^m$	$\mathbb{R}^{n\times m}$	$\frac{\partial f_i}{\partial x_j}$	Jacobian
$\mathbb{R}^{m\times n}$	$\mathbb{R}^{m\times n}$	$\mathbb{R}^{m\times n}$	$\frac{\partial f}{\partial x_{ij}}$	元素导数

微分记号的常用性质 在矩阵分析和优化中，微分记号d有如下常用性质：

$$\begin{aligned}
dA &= 0 \\
d(\alpha X) &= \alpha\, dX \\
d(AXB) &= A\, dX\, B \\
d(X+Y) &= dX + dY \\
d(X^\top) &= (dX)^\top \\
d(XY) &= (dX)Y + X(dY) \\
d\langle X, Y\rangle &= \langle dX, Y\rangle + \langle X, dY\rangle \\
d\left(\frac{X}{\phi}\right) &= \frac{\phi\, dX - (d\phi)X}{\phi^2} \\
d\,\mathrm{tr}\, X &= I \\
df(g(x)) &= \frac{df}{dg}\cdot dg(x)
\end{aligned}$$

这些性质在推导链式法则、矩阵函数的导数等问题时非常有用。

下面给出例子，使用逐元素求导和使用公式求导：

例子 3.6 (逐元素求导)**.** 已知 $f(x) = \frac{1}{2}x^T Ax + b^T x + c$，其中 $A \in \mathbb{R}^{n\times n}$，$b \in \mathbb{R}^n$，$c \in \mathbb{R}$。求 $\nabla f(x)$。

解： 展开 $f(x)$ 得

$$f(x) = \frac{1}{2}\sum_{i,j=1}^{n} x_i a_{ij} x_j + \sum_{i=1}^{n} b_i x_i + c$$

对 x_i 求偏导：

$$\frac{\partial f(x)}{\partial x_i} = \frac{1}{2}\sum_{j=1}^{n}(a_{ij} + a_{ji})x_j + b_i$$

因此

$$\nabla f(x) = \frac{1}{2}(A + A^T)x + b$$

例子 3.7 (利用公式求导)**.** 已知 $f(X) = \mathrm{tr}(AX)$，其中 $A, X \in \mathbb{R}^{n\times n}$。求 $\nabla f(X)$。

解： 利用矩阵求导公式 $\nabla_X \mathrm{tr}(AX) = A^T$，所以

$$\nabla f(X) = A^T$$

3.4 凸函数及其性质

凸函数是满足Jesen不等式的函数，即

$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta)f(y), \quad \forall x, y, \theta \in [0,1] \tag{2}$$

即对函数任意两点连起来，两点之间的函数值不超过两点函数值的线性组合。如果严格取不等号则为strictly convex function（严格凸函数）

对于不等式方向，又有concave function（凹函数）。

3.4.1 凸函数的上方图（Epigraph）

定义 3.8 (上方图（Epigraph）). 给定函数 $f : \mathbb{R}^n \to \mathbb{R}$，其上方图定义为

$$\text{epi}(f) = \{(x, y) \mid f(x) \le y\}$$

即所有在图像上方（含图像本身）的点的集合。

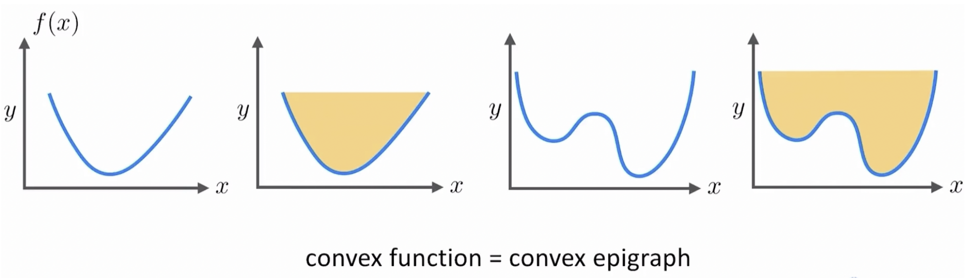

图 2: 凸函数的上方图（Epigraph）。左：凸函数的上方图是凸集；右：非凸函数的上方图不是凸集。

性质

- 函数 f 是凸函数，当且仅当其上方图 epi(f) 是凸集。
- 直观理解：凸函数的图像“开口向上”，其上方图不会出现“凹陷”。

凸函数的sub level set（下水平集）定义为

$$\text{sub}(f) = \{x \mid f(x) \le c\}$$

是凸集当且仅当 f 是凸函数。

具有凸的下水平集的函数不一定是凸函数，比如$f(x) = log(|x| + .1)$，这是quasi-convex函数，但不是凸函数。quasi-convex函数的下水平集是凸集，但上方图不一定是凸集。

凸的性质很容易被保留下来，而quasi-convex的性质不容易被保留下来。

凸函数的局部最小值也是全局最小值，这就是为什么我们要研究凸函数的原因。

例： $f(x) = x + x^2 + x^3 + |x|$是凸集。

各种范数也是凸的，比如 L_1 范数、L_2 范数等。范数满足三角不等式，即$\|x+y\| \le \|x\| + \|y\|$，则有$\theta a + (1-\theta)b \le \theta\|a\| + (1-\theta)\|b\|$，因此范数是凸函数。

纺射变换也是凸的，即若 $f(x)$ 是凸函数，则 $g(x) = f(Ax + b)$ 也是凸函数，其中 A 是线性变换，b 是平移向量。

点对点最大值操作与凸性保持 点对点最大值操作（point-wise max）可以保持凸性。设有一组凸函数 $f_i(x)$，定义

$$g(x) = \max_i f_i(x)$$

则 $g(x)$ 也是凸函数。

常见例子

- **绝对值**：$|x| = \max\{x, -x\}$，是两个线性函数的最大值，因此是凸函数。
- **无穷范数**：$\|x\|_\infty = \max_i |x_i|$，是若干个凸函数（$|x_i|$）的最大值，因此是凸函数。
- **最大特征值**：$\|A\|_2 = \max_v v^T Av$，是关于 A 的凸函数（A 为对称矩阵时）。

上方图的凸性是否保持？ 点对点最大值操作不仅保持函数的凸性，也保持上方图（epigraph）的凸性。因为若 $f_i(x)$ 的上方图都是凸集，则

$$\mathrm{epi}(g) = \bigcap_i \mathrm{epi}(f_i)$$

交集仍为凸集，因此 $g(x)$ 的上方图也是凸集。

注释 3.9. 点对点最大值操作是构造新凸函数的常用方法，常见于范数、绝对值、最大特征值等场景。

3.5 为什么这些函数是凸的？

- **Trace**：$f(X) = \mathrm{trace}(A^T X)$ 是线性算子，线性函数总是凸的（也是凹的）。
- **Distance over set**：$f(x) = \max_{y\in C} \|x - y\|$，对凸集 C，$\|x - y\|$ 关于 x 是凸函数，对 y 取最大值仍保持凸性（最大值操作保持凸性）。
- **Distance to convex set**：$f(x) = \min_{y\in C} \|x - y\|$，这是特殊情形：$C$ 为凸集时，$\|x - y\|$ 关于 x 是凸函数，对 y 取最小值，结果也是凸函数。
- **Affine Norm**：$f(x) = \|b + \sum_i A_i x_i\|_2$，仿射变换加范数，范数是凸函数，仿射变换不会破坏凸性，因此整体是凸函数。

一般结论 如果 $g(x, y)$ 关于 x 是凸函数，则对 y 取最大值或最小值（在 y 上凸/凹时）通常能保持凸性。即

$$f(x) = \sup_{y\in C} g(x, y)$$

若 $g(x, y)$ 关于 x 是凸函数，则 $f(x)$ 也是凸函数。

注释 3.10. 这些性质在优化建模中非常重要，常用于构造新的凸目标函数或约束。

对于凸函数，

凸函数的切线性质　凸函数在任意点的切线（线性近似）都在函数图像的下方，即

$$f(y) \geq f(x) + \nabla f(x)^T (y - x), \quad \forall x, y \in \mathrm{dom}\, f$$

这意味着：**凸函数总是位于其线性近似的上方。**

该性质是凸优化理论的基础，保证了梯度下降等方法的收敛性和全局最优性。

则有：

一阶最优性条件与全局最优性　对于凸函数，若 x^* 是可微函数 $f(x)$ 的极小点，则有

$$\nabla f(x^*) = 0$$

结合凸函数的切线性质，得到

$$f(y) \geq f(x^*) + \nabla f(x^*)^T (y - x^*) = f(x^*)$$

即对于所有 y，$f(y) \geq f(x^*)$，因此 x^* 是全局最小点。

注释 3.11. 对于非凸函数，$\nabla f(x^*) = 0$ 只保证 x^* 是驻点，可能是鞍点、局部极小或极大值点。只有凸函数才保证一阶最优性条件对应全局最优。

结论　**对于凸函数，任何极小点都是全局极小点。**

3.5.1　二阶条件（Second-order conditions）

凸函数的二阶判据　一个光滑函数 $f(x)$ 是凸函数，当且仅当其Hessian矩阵处处半正定：

$$\nabla^2 f(x) \succeq 0, \quad \forall x$$

即Hessian为半正定矩阵（semi-definite）。

非凸函数的极小值二阶条件　对于非凸函数，极小点 x^* 满足：

$$\nabla^2 f(x^*) \succeq 0$$

即在极小点处Hessian半正定。

Hessian的作用　Hessian矩阵是光滑函数的良好局部二次近似模型。二阶条件在优化理论和算法中非常重要。

3.5.2　强凸性（Strong Convexity）

定义　函数 f 是 m-强凸（m-strongly convex）的，如果存在 $m > 0$，使得对任意 x, y，有

$$f(y) \geq f(x) + (y - x)^T \nabla f(x) + \frac{m}{2} \|y - x\|^2$$

其中 m 被称为最小曲率（min curvature），该条件对任意凸函数都成立。

Hessian 存在时的等价条件　当 f 二阶可微时，泰勒展开有

$$f(y) \approx f(x) + (y-x)^T \nabla f(x) + \frac{1}{2}(y-x)^T \nabla^2 f(x)(y-x)$$

强凸性要求

$$f(y) \geq f(x) + (y-x)^T \nabla f(x) + \frac{\lambda_{\min}}{2}\|y-x\|^2$$

其中 $\lambda_{\min}$ 是 Hessian $\nabla^2 f(x)$ 的最小特征值。

矩阵判据　因此，f 是 m-强凸的当且仅当

$$\nabla^2 f(x) \succeq mI$$

即 Hessian 处处大于等于 m 倍单位阵。

注释 3.12. 强凸函数不仅保证唯一极小点，还能保证优化算法更快收敛（如梯度下降线性收敛）。

强凸用于判别函数的收敛性和唯一性。强凸函数的梯度下降法具有更好的收敛速度。强凸这里构建了一个二次函数的下界，保证了函数的“弯曲程度”，使得优化问题更易处理。

这里构建上界：

3.5.3 Lipschitz 连续与上界（Lipschitz Smoothness and Upper Bound）

定义　若函数 f 的梯度 $\nabla f(x)$ 是 M-Lipschitz 连续的，即存在常数 $M > 0$，使得

$$\|\nabla f(x) - \nabla f(y)\| \leq M\|x-y\|, \quad \forall x, y$$

则称 f 是 M-smooth 的，M 称为 Lipschitz 常数。

二次上界　M-smooth 的函数 f 满足如下二次上界：

$$f(y) \leq f(x) + (y-x)^T \nabla f(x) + \frac{M}{2}\|y-x\|^2$$

即函数值不会超过其在 x 点的线性近似加上一个二次项。

推论　该上界在分析梯度下降等优化算法时非常重要，常用于收敛性证明和步长选择。

目标误差与最优解距离的界　对于 M-smooth、m-强凸的函数，有如下界：

$$\frac{m}{2}\|y-x^*\|^2 \leq f(y) - f(x^*) \leq \frac{M}{2}\|y-x^*\|^2$$

其中 x^* 为最小点。这说明目标函数值与最优解的距离平方成正比。

3.5.4 凸函数的条件数（Condition Number）

定义　条件数 κ 衡量函数的“弯曲程度”或“各向异性”，反映优化问题的难易程度。常见定义如下：

- **一般函数**：$\kappa = \dfrac{\text{major axis}}{\text{minor axis}}$，即等高线主轴与次轴之比。
- **光滑函数**：$\kappa \approx \text{cond}(\nabla^2 f(x))$，即 Hessian 的条件数（最大特征值与最小特征值之比）。
- **可微函数**：$\kappa = \dfrac{M}{m}$，其中 M 是 Hessian 最大特征值，m 是最小特征值。

几何意义　条件数越大，等高线越"扁"，优化算法收敛越慢；条件数越小，等高线越"圆"，优化更容易。

优化算法中的作用　条件数决定了梯度下降等一阶方法的收敛速度。条件数越大，收敛越慢，此时需要考虑更高阶的信息；条件数越小，收敛越快。实际问题中常通过预处理（如预条件化）降低条件数以加速优化。

3.5.5　次微分 sub-differential

对于不可微的凸函数，梯度的概念可以推广为**次微分**。在点 x 处，$f(x)$ 的次微分定义为：

$$\partial f(x) = \left\{ g \mid f(y) \geq f(x) + (y-x)^T g, \forall y \right\}$$

即所有使得该不等式成立的向量 g 构成次微分集。

几何意义　在可微点，次微分只有一个元素，即梯度本身；在不可微点（如 $|x|$ 在 $x=0$ 处），次微分是所有支持该点的切线斜率的集合。

最优性条件　对于凸函数，x^* 是极小点当且仅当

$$0 \in \partial f(x^*)$$

即 0 属于 x^* 处的次微分集。

例子 3.13 (绝对值函数的次微分). 设 $f(x) = |x|$，则

$$\partial f(x) = \begin{cases} \{1\}, & x > 0 \\ \{-1\}, & x < 0 \\ [-1, 1], & x = 0 \end{cases}$$

即在 $x = 0$ 处，所有斜率在 $[-1, 1]$ 区间内的直线都是 $|x|$ 的支持切线。

例子 3.14 (最大值函数的次微分). 设 $f(x) = \max\{x_1, x_2, \ldots, x_n\}$，则在 x 处的次微分为

$$\partial f(x) = \text{conv}\left\{ e_i \mid x_i = f(x) \right\}$$

其中 e_i 是第 i 个标准基向量，conv 表示取凸包。即所有达到最大值的分量对应的基向量的凸组合。

沿着sub-differential的方向前进，函数值不会下降。次微分的概念在凸优化中非常重要，尤其是处理不可微凸函数时。

对于非光滑函数，最速下降的方向是次微分集中的膜长最小的元素的反方向。

3.5.6 凸函数的单调性与一阶条件

（次）梯度的单调性（Monotonicity） 任何凸函数的（次）梯度是单调的，即对于任意 x, y，有

$$\langle y - x, \nabla f(y) - \nabla f(x) \rangle \geq 0$$

，这里是内积。更一般地，对于不可微的凸函数，若 $g_x \in \partial f(x),\ g_y \in \partial f(y)$，则

$$\langle y - x, g_y - g_x \rangle \geq 0$$

推导 该性质可以通过将凸函数的切线性质在 x 和 y 处分别写出并相加得到：

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$
$$f(x) \geq f(y) + \nabla f(y)^T (x - y)$$

两式相加，得

$$0 \geq (\nabla f(x) - \nabla f(y))^T (y - x)$$

即

$$\langle y - x, \nabla f(y) - \nabla f(x) \rangle \geq 0$$

意义 该单调性是凸优化理论的基础之一，保证了梯度型方法的收敛性，也可用于证明极值点的唯一性等性质。

3.6 无约束的非凸优化

3.6.1 最速下降 steepest gradient descent

$$x^{k+1} = x^k - \eta \nabla f(x^k) \tag{3}$$

3.6.2 步长选择与线搜索（Step Size and Line Search）

最速下降法的关键在于步长（step size, 记作 η）的选择。常见的步长选择策略包括：

- **常数步长（Constant step size）**：$\eta = c$，其中 c 为常数。
- **递减步长（Diminishing step size）**：$\eta = c/k$，k 为迭代次数。
- **精确线搜索（Exact line search）**：$\eta = \arg\min_\alpha f(x^k + \alpha d)$，在当前方向 d 上精确最小化目标函数，即在当前状态下使系统下降最多的步长。
- **非精确线搜索（Inexact line search）**：η 满足

  $$\eta \in \left\{ \alpha \mid f(x^k) - f(x^k + \alpha d) \geq -c \cdot \alpha d^\top \nabla f(x^k) \right\}$$

  其中 $c \in (0, 1)$ 是常数，保证每步都有足够的下降。

不同的步长策略影响算法的收敛速度和稳定性。精确线搜索通常收敛较快，但计算代价高；常数步长实现简单，但需手动调参；非精确线搜索（如Armijo规则）在实际中应用广泛，兼顾效率与收敛性。

3.6.3 Armijo条件（充分下降条件）

在非精确线搜索中，常用的步长选择准则之一是**Armijo条件**（sufficient decrease condition），其数学表达为：

$$\eta \in \left\{\alpha \mid f(x^k) - f(x^k + \alpha d) \geq -c \cdot \alpha d^\top \nabla f(x^k)\right\}, \quad c \in (0, 1)$$

即步长α需要使目标函数的下降量至少达到梯度下降的一个比例。

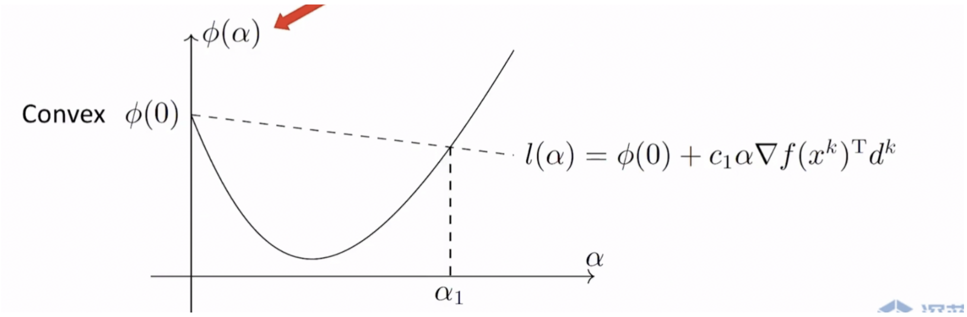

图 3: Armijo条件的几何解释：$\phi(\alpha)$为目标函数在方向d上的截面，$l(\alpha)$为线性下界。

如上图所示，$\phi(\alpha)$表示在当前点沿方向d的目标函数截面，$l(\alpha) = \phi(0)+c_1\alpha\nabla f(x^k)^\top d^k$为线性下界。Armijo条件要求实际下降量不小于该线性下界。

意义

3.6.4 Backtracking/Armijo线搜索算法

Algorithm 1: Backtracking/Armijo线搜索

Input: 当前点 x^k，目标函数 $f(x)$，梯度 $\nabla f(x^k)$，初始步长 $\eta > 0$，常数 $c \in (0, 1)$

Output: 步长 η，新点 x^{k+1}

选择搜索方向 $d = -\nabla f(x^k)$;

while $f(x^k + \eta d) > f(x^k) + c \cdot \eta d^T \nabla f(x^k)$ **do**

  $\eta \leftarrow \eta/2$;

更新：$x^{k+1} = x^k + \eta d$;

重复上述过程，直到梯度足够小或次微分包含零为止。我们在最优控制中也用到了这个下降方法。

结论：在迭代精度要求更高的时候性能是相似的。最速梯度下降因为只是用了一阶信息，所以收敛速度较慢，特别当条件数较大时，收敛速度会很慢。

3.7 modified damped newton method （修正阻尼牛顿法）

需要函数连续，一阶和二阶导数连续。

3.7.1 Newton法与二阶泰勒展开

二阶泰勒展开 对光滑函数 $f(x)$，在点 x_k 处的二阶泰勒展开为

$$f(x) \approx \hat{f}(x) \triangleq f(x_k) + \nabla f(x_k)^T(x - x_k) + \frac{1}{2}(x - x_k)^T\nabla^2 f(x_k)(x - x_k)$$

最小化二次近似　对二次近似 $\hat{f}(x)$ 求极小值，令梯度为零：

$$\nabla \hat{f}(x) = \nabla^2 f(x_k)(x - x_k) + \nabla f(x_k) = 0$$

解得

$$x = x_k - \left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k)$$

前提是 $\nabla^2 f(x_k) \succ 0$（正定）。

Newton步　因此，Newton法的迭代公式为

$$\boxed{x_{k+1} = x_k - \left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k)}$$

如果 f 是二次函数，则Newton法一步即可到达最优解。

注释 3.15. Newton法利用了二阶信息（Hessian），在Hessian正定时具有二次收敛速度，但每步需解线性方程组，计算代价较高。实际中常结合阻尼或修正策略以保证收敛性。

3.7.2 Hessian修正与阻尼牛顿法（Modified/Damped Newton Method）

Hessian矩阵不一定一直可逆，当其不可逆的时候，Newton法就不适用了。此时需要修正Hessian矩阵，使其可逆，我们已经在最优控制中学习过一遍了，这里稍加复习：

Hessian修正　当Hessian矩阵$\nabla^2 f(x_k)$不可逆或不正定时，常用的修正方法有：

- **加正则项（Levenberg-Marquardt修正）**：在Hessian上加上一个正定矩阵（通常是λI），即

  $$H_{\text{mod}} = \nabla^2 f(x_k) + \lambda I$$

  其中$\lambda > 0$，保证H_{mod}正定且可逆。

- **截断负特征值**：将Hessian的负特征值调整为正值，确保正定性。

- **BFGS等拟牛顿方法**：用正定近似矩阵代替Hessian。

阻尼牛顿法　为防止步长过大导致发散，采用阻尼（damping）策略，即

$$x_{k+1} = x_k - \eta \left[H_{\text{mod}}\right]^{-1} \nabla f(x_k)$$

其中$\eta \in (0, 1]$为阻尼系数，通常结合线搜索（如Armijo条件）自适应调整。

算法流程

1. 计算梯度$\nabla f(x_k)$和Hessian$\nabla^2 f(x_k)$。
2. 若Hessian不正定或不可逆，则加正则项λI修正。
3. 求解修正后的牛顿步$d_k = -\left[H_{\text{mod}}\right]^{-1} \nabla f(x_k)$。
4. 用线搜索确定步长η，更新$x_{k+1} = x_k + \eta d_k$。
5. 重复直到收敛。

注释 3.16. 修正和阻尼策略保证了牛顿法在非凸或病态问题下的收敛性和稳定性，是实际优化中常用的改进方法。

3.8 Practical Newton's Method 实用牛顿法

基本思想 Newton法的核心是解线性方程组

$$\nabla^2 f(x) d = -\nabla f(x)$$

但实际中Hessian矩阵可能是半正定（PSD）或不定（indefinite），导致数值不稳定或不可逆。

凸函数情形 若f是凸函数，Hessian必为半正定。此时可采用如下修正：

$$M = \nabla^2 f(x) + \epsilon I, \quad \epsilon = \min\left(1, \|\nabla f(x)\|_\infty\right)/10$$

其中ϵ为小正数，I为单位阵，确保M正定。

Cholesky分解 由于M正定，可用Cholesky分解$M = LL^\top$，L为下三角阵。搜索方向d由

$$Md = -\nabla f(x)$$

求解。

非凸函数情形 若f非凸，Hessian可能不定。此时可用Bunch-Kaufman分解：

$$M = LBL^\top$$

其中B为块对角阵（块大小1×1或2×2），所有2×2块包含非正特征值，可方便地修正。

小结

- 对于凸问题，Hessian加正则项后用Cholesky分解，保证数值稳定。
- 对于非凸问题，采用Bunch-Kaufman分解，动态修正Hessian的非正定性。
- 这两种方法都能有效提升Newton法在实际优化中的鲁棒性和收敛性。