# **Maths**

# 雅可比矩阵 Jacobi Matrix

差分方程系统?

# 黑塞矩阵 Hessian Matrix

# 仿射函数 Affine Function

形如 f(x) = a + bx。对应的,严格意义上的线性函数是 f(x) = bx, a = 0

### 性质

- 1. 当且仅当  $f''(x) \equiv 0$
- 2. 当且仅当对  $\forall \ 0 \le p \le 1, f[pu + (1-p)u] \equiv pf(u) + (1-p)f(v)$ ,即凸组合直接可分

# 位似函数 Homothetic Function

(一次) 齐次函数的单调变换

 $f(\mathbf{x})$  是一个位似函数当且仅当它可以表示成  $f(\mathbf{x})=g(h(\mathbf{x}))$ ,其中  $h(\cdot)$  是(一次)齐次函数,  $g(\cdot)$  是(正的)单调函数

### 欧拉法则

如果 f 是一个可微的一次齐次函数,那么

$$f(\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial f(\mathbf{x})}{\partial x_i} x_i \tag{1}$$

证明为对恒等式  $f(t\mathbf{x}) \equiv t f(\mathbf{x})$  两边对 t 求微分,然后令 t=1

# 分析

**紧集**:特殊情况下简单理解  $S \subset \mathbb{R}^n$  是有界的闭集,那么 S 是紧集。

注意有/无界和集合开闭并不存在——对应关系,也就是可以存在无界的闭空间。从度量? 点集拓扑领域? 角度理解?

### 两个例子:

- 如果全集选取为有理数集,则 ℝ 既是开集也是闭集,其补集 ∅ 空集也是既开又闭
- ullet 如果全集选取为  $R^n$ ,则实数轴应该为闭集;如果全集就是实数轴,那实数轴应该既开又闭

## 矩阵定性

### 正定

- ⇔特征值全部大于零
- ⇔ 顺序主子式全为正
- $\Leftrightarrow$  存在实可逆矩阵 C ,分解为  $A=C^TC$  ;就是和 E 合同
- $\Leftrightarrow$  存在秩为 n 的  $m \times n$  实矩阵 B,使  $A = B^T B$
- $\Leftrightarrow$  存在主对角元素全为正的实三角矩阵 R,使  $A=R^TR$
- ⇒ 逆矩阵也正定
- ⇒对角线元素均为正数

### 负定

- ⇔ 特征值全部小于零
- $\Leftrightarrow$  顺序主子式的符号为  $(-1)^k$
- ⇒ 逆矩阵也负定
- ⇒对角线元素均为负数

### 半定

对所有  $\mathbf{x}$  都有  $\mathbf{x}^T A \mathbf{x} \begin{pmatrix} \leq \\ > \end{pmatrix} 0$ ,即允许当  $\mathbf{x} \neq \mathbf{0}$  时  $\mathbf{x}^T A \mathbf{x} = 0$  仍成立,放松了要求

### 半正定

- ⇔顺序主子式非负
- ⇔ 特征值全部非负
- $\Leftrightarrow$  存在实矩阵 C ,分解为  $A = C^T C$

### 半负定

### 约束条件下的定性

如果不要求对  $\forall \mathbf{x}$ , $\mathbf{x}^T A \mathbf{x}$  都要有确定的符号,而只是要求某些限定的  $\mathbf{x}$  成立,比如说满足  $\mathbf{b} \mathbf{x} = 0$  且  $\mathbf{x} \neq \mathbf{0}$  的  $\mathbf{x}$ 。我们就称为 A 在约束条件下的定性。具体可以用 **加边黑塞矩阵(增广黑塞矩阵)** 来进行判断

$$\begin{pmatrix} 0 & b_1 & \cdots & b_n \\ b_1 & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_n & a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

$$(2)$$

- 约束下正定, 当且仅当加边黑塞矩阵主子式均为负 (**注意这和原来的单纯正定刚好相反**)
- 约束下负定,当且仅当加边黑塞矩阵主子式符号为  $(-1)^k$ , $k=2,3,\cdots$  n (**这个却和单纯负定一致**)

这个加边矩阵的来历是基于约束下的拉格朗日方法

特别地,如果黑塞矩阵在约束下是负定的(不止是半负定),那么我们称函数有 **正则最大值**(Regular Maximum)。一个正则最大值肯定是一个严格的局部最大值,但反之不一定

## 梯度 Gradient

考虑一个线性泛函  $f: \mathbb{R}^n \to \mathbb{R}, \ f$  在  $\mathbf{x}^*$  的梯度是一个向量,其坐标就是该处的各个偏导数

$$\mathbf{D}f(\mathbf{x}^*) = \left(\frac{\partial f(\mathbf{x}^*)}{\partial x_1}, \dots, \frac{\partial f(\mathbf{x}^*)}{\partial x_n}\right)$$
(3)

注意梯度有其方向,并且指向 f 增加最快的方向。证明为取一个单位向量  $\mathbf{h}$ , f 在  $\mathbf{x}^*$  处沿  $\mathbf{h}$  方向的导数即为  $\mathbf{D}f(\mathbf{x}^*)\mathbf{h}$  。内积为  $\mathbf{D}f(\mathbf{x}^*)\mathbf{h} = |\mathbf{D}f(\mathbf{x}^*)|cos\theta$ 。显然当  $\theta=0$  时亦即两向量共线时取到最大值

# 超平面 Hyperplane

对于法向量  $\omega$ ,超平面公式可以写为  $\omega^T x + b = 0$ 

# 切平面 Tangent Hyperplane

$$H(a) = \{\mathbf{x} : \mathbf{D}f(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) = 0\}$$

$$\tag{4}$$

就是泰勒展开第二项的线性近似

如果  $\mathbf{x}$  是切超平面中的一个向量,那么  $\mathbf{x} - \mathbf{x}^*$  与 f 在  $\mathbf{x}^*$  处的梯度是正交的

## 分离超平面

如果 A 和 B 是  $R^n$  中两个非空不相交的凸集,那么就可以从中插入一个超平面。存在一个线性泛函  $\mathbf{p}$  ,使得所有 A 中的  $\mathbf{x}$  和 B 中的  $\mathbf{y}$  都有  $\mathbf{p}\mathbf{x} \geq \mathbf{p}\mathbf{y}$  成立

# 凸组合

向量  $\mathbf{u}$  和  $\mathbf{v}$  的线性组合

$$\theta \mathbf{u} + (1 - \theta) \mathbf{v}, \qquad (0 \le \theta \le 1)$$
 (5)

可以解释为两个向量的加权平均

# 凸集 Convex Set

#### 定义

几何上可以这么理解,要求点集无孔,边缘各处无缩进

 $A \in \mathbb{R}^n$  中的一个点集,如果  $\mathbf{x}$  和  $\mathbf{y}$  都在 A 中,且他们的凸组合仍然在 A 中,则 A 是凸集。如果对于 0 < t < 1 的所有 t 都有  $t\mathbf{x} + (1 - t)\mathbf{y}$  在 A 的内部(取消了等于号),则称 A 为严格凸的

#### 性质

凸集的和仍然是凸的

## 凸函数 Convex Function

定义

1. 几何直观定义: 凸组合线段的高度大于等于凸组合函数值的高度; 或者说总在切平面以上(或与之重合)

对于函数 f 定义域内任意两个不同的点 u 和 v,且相对于  $0 < \theta < 1$ ,当且仅当

$$\theta f(u) + (1 - \theta)f(v) \ge f[\theta u + (1 - \theta)v] \tag{6}$$

时,f为凸函数

2. 函数可微时的定义:

若对于定义域内任意给定点  $u=(u_1,\cdots,u_n)$  和另一给定点  $v=(v_1,\cdots,v_n)$ ,当且仅当

$$f(v) \ge f(u) + \sum_{j=1}^{n} f_j(u)(v_j - u_j)$$
 (7)

可微函数 f 为凸函数。其中  $f_i(u) \equiv \partial f/\partial x_i$  在  $u = (u_1, \dots, u_n)$  计算其值

3. 去掉等号即为严格凸

### 性质

- 1. 函数的和
  - 1. 多个凸函数的和仍然是凸函数
  - 2. 只要其中一个凸函数是严格凸的,那么多个凸函数的和就是严格凸的
- 2. 更大的拟凸的性质, 拟凸函数可以引致一个凸集

$$S^{\leq} = \{x | f(x) \leq k\}, \qquad [f(x)$$
 为凸函数,  $k$  为任意常数 [

此下轮廓集为一个凸集

3. 严格凸函数的黑塞矩阵(一般)是正定的;凸函数的黑塞矩阵是半正定的。可以就此判别

# "凸"的语义差别

● 在集合中: 凸描述的是集合中的点如何"填充"到一起。强调的是"内外之分"

● 在函数中: 凸描述的是一条曲线或者曲面如何弯曲。强调的是"上下之分"

# 拟凸函数 QuasiConvex Function

### 定义

1. 直观定义

对于  $\forall a \in (0,1)$ ,都有  $f(ax+(1-a)y) \leq max\{f(x),f(y)\}$  成立。拟凹就是大于等于最小值也就是说函数 f 图形上任意弧段 MN (设点 N 高于或等于 M),除了两个端点以外这个弧段上任意一点高度都低于或等于点 N 的高度

2. 轮廓集定义

对于任意常数 k, 当且仅当

$$S^{\leq} \equiv \{x | f(x) \leq k\} \tag{9}$$

此下轮廓集是凸集时,函数 f(x) 是拟凸的。变量也可以是一个向量  ${f x}$  。拟凹就是上轮廓集为凸集

#### 性质

- 1. 凸 ⊂ 拟凸
- 2. 同时有凹凸区段的函数也可以是整体拟凹的,典型比如正态分布密度函数

### 判别

拟凹的矩阵条件

$$\begin{pmatrix} 0 & f_1 & f_2 & \cdots & f_n \\ f_1 & f_{11} & f_{12} & \cdots & f_{1n} \\ f_2 & f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & & & & & \\ f_n & f_{n1} & f_{n2} & \cdots & f_{nn} \end{pmatrix}$$

$$(10)$$

注意此矩阵与矩阵定性部分约束条件下的判定矩阵不同,加边部分是自己的一阶导数而不是约束向量。 此矩阵的顺序主子式的符号为 $(-1)^k$ 时,f为拟凹

# 包络定理 Envelope Theorem

值函数(Value Function)为 M(a)=f(x(a),a),其中自变量 x 根据参数 a 取得最优化。想要考察的是最优值如何对参数 a 变化作出反应,也就是  $\frac{\partial M}{\partial a}$ 

值函数两边对a求导

$$\frac{dM(a)}{da} = \frac{\partial f[x(a), a]}{\partial x} \frac{\partial x(a)}{\partial a} + \frac{\partial f[x(a), a]}{\partial a}$$
(11)

而由于x(a)已是最优化选择,所以一阶条件有

$$\frac{\partial f[x(a), a]}{\partial x} = 0 \tag{12}$$

从而原式只剩下

$$\frac{dM(a)}{da} = \frac{\partial f[x, a]}{\partial a}|_{x=x(a)}$$
(13)

也就是说,**值函数关于参数的全导数,等于在最优选择计算的偏导数**。参数 a 变化时有两个效应

- 直接影响 f
- 通过 x 间接影响 f

但因为最优化,间接效应等于零,只剩下直接效应

### 带有约束的情况

考虑如下值函数

$$M(a) = \max_{x_1, x_2} g(x_1, x_2, a)$$
 (14)  
 $s.t. \quad h(x_1, x_2, a) = 0$ 

注意到参数 a 可以同时出现在值函数和约束条件之中

对应拉格朗日函数  $L = g(x_1, x_2, a) - \lambda h(x_1, x_2, a)$ , 一阶条件为

$$\frac{\partial g}{\partial x_1} - \lambda \frac{\partial h}{\partial x_1} = 0$$

$$\frac{\partial g}{\partial x_2} - \lambda \frac{\partial h}{\partial x_2} = 0$$

$$h(x_1, x_2, a) = 0$$
(15)

对值函数  $M(a) \equiv g(x_1(a), x_2(a), a)$  两边求导

$$\frac{dM}{da} = \frac{\partial g}{\partial x_1} \frac{dx_1}{da} + \frac{\partial g}{\partial x_2} \frac{dx_2}{da} + \frac{\partial g}{\partial a}$$
(16)

通过上一步的结果替换前两项

$$\frac{dM}{da} = \lambda \left(\frac{\partial h}{\partial x_1} \frac{dx_1}{da} + \frac{\partial h}{\partial x_2} \frac{dx_2}{da}\right) + \frac{\partial g}{\partial a} \tag{17}$$

而由于最优化,满足约束  $h(x_1(a), x_2(a), a) \equiv 0$ ,两边对 a 求微分

$$\frac{\partial h}{\partial x_1} \frac{dx_1}{da} + \frac{\partial h}{\partial x_2} \frac{dx_2}{da} + \frac{\partial h}{\partial a} = 0 \tag{18}$$

代入上式即可得到

$$\frac{dM}{da} = -\lambda \frac{\partial h}{\partial a}|_{x=x(a)} + \frac{\partial g}{\partial a}|_{x=x(a)}$$
(19)

总结来说也就和无约束情况类似,想要求值函数关于参数的变动情况,**直接将拉格朗日函数对该参数求导,并代入自变量的最优化取值**。同样只有直接效应没有间接效应

两个引理可以直接记忆为用求导代替除法,包络的思路是从等号右边推出左边,即从值函数最优化推出 自变量如何跟随参数变动取得最优化

## 霍特林引理 Hotelling's lemma

通过利润函数求出净供给函数,已知 $\pi(\mathbf{p}) = \mathbf{p}\mathbf{y}(\mathbf{p})$ 

应用无约束包络定理

$$y_i(\mathbf{p}) = \frac{\partial \pi(\mathbf{p})}{\partial p_i}|_{\mathbf{p} = \mathbf{p}^*} \quad i = 1, \dots, n$$
 (20)

## 谢泼德引理 Shephard's lemma

通过成本函数求出要素需求函数,已知  $c(\mathbf{w}, y)$ ,约束条件  $f[\mathbf{x}(\mathbf{w}, y)] \equiv y$ 

应用有约束包络定理

$$x_i(\mathbf{w}, y) = \frac{\partial C(\mathbf{w}, y)}{\partial w_i} \quad i = 1, \dots, n$$
 (21)

## 库恩-塔克条件 Kuhn-Tucker conditions

和 KKT 条件(Karush-Kuhn-Tucker conditions)是一回事。一般可粗略认为是求最值的**必要条件**,有 多个极值解时要再比较

具体条件可以记忆为

$$egin{aligned} \max_{x} \ f(x) \ s. \ t. \ h_{j}(x) = 0, \ j = 1, 2, \cdots, q \ g_{i}(x) \leq 0, \ i = 1, 2, \cdots, p \end{aligned}$$

则

$$\nabla f(x^*) = \sum_{j} \lambda_j \nabla h_j(x^*) + \sum_{i} \mu_i \nabla g_i(x^*)$$

$$\mu_i \ge 0, \mu_i g_i(x^*) = 0$$
(23)

也就是说在极值处,目标函数的梯度(函数值增长最快的方向)可以表示为约束条件梯度的线性组合。 注意这一条件移项后才是拉格朗日函数求梯度后的形式,这会导致拉格朗日乘子的符号相反!

以n个变量和m个不等式约束条件为例:

$$max \quad \pi = f(x_1, \dots, x_n)$$

$$s.t. \quad g^i(x_1, \dots, x_n) \le r_i$$
(24)

第一步写出纯经典拉格朗日函数

$$Z = f(x_1, \dots, x_n) - \sum_{i=1}^{m} \lambda_i [g^i(x_1, \dots, x_n) - r_i]$$
 (25)

注意  $\lambda$  后跟的约束条件符号方向最好保证和约束条件中的  $g^i(\mathbf{X})$  以及不等号方向一致, $\lambda$  前一律设为负号,方便与原条件统一记忆

- 确保对 λ 求偏导后写出的不等式条件符合原式约束要求
- $\lambda$  的符号: 这里是求约束条件下的**最大值**,所以边界解时 f 的梯度  $\nabla f$  应该指向可行域  $q^i(\mathbf{X}) < r_i$  的**外部**(因为**可行域为小于等于号**),也就是说和  $\nabla g$  **同向**

库恩塔克条件具体为:

最大化条件 
$$\frac{\partial Z}{\partial x_j}=0$$
 (26) 复述约束条件  $\frac{\partial Z}{\partial \lambda_j}\geq 0, \quad$  拉格朗日乘数约束 $\lambda_i\geq 0, \quad$  且 互补松弛条件 $\lambda_i\frac{\partial Z}{\partial \lambda_j}=0$  这里 $\lambda_i$ 非负,表明 $\nabla f$ 与 $\nabla g$ 同向

如若换为最小化问题

$$min \quad \pi = f(x_1, \dots, x_n)$$

$$s.t. \quad g^i(x_1, \dots, x_n) \le r_i$$

$$(27)$$

**拉格朗日函数不变**,当然可以用对偶化为  $max - \pi$ ,也可直接记忆为

最小化条件 
$$\dfrac{\partial Z}{\partial x_j}=0$$
 (28) 复述约束条件  $\dfrac{\partial Z}{\partial \lambda_j}\geq 0$ , 拉格朗日乘数约束 $\lambda_i\leq 0$ , 且 互补松弛条件 $\lambda_i\dfrac{\partial Z}{\partial \lambda_j}=0$  这里 $\lambda_i$ 非正,因为求约束条件下最小值,边界解时 $f$ 的梯度 $\nabla f$ 应该指向可行域 $g^i(X)\leq r_i$ 的内部 因为可行域为小于等于号,也就是说和 $\nabla g$ 反向

**互补松弛**:对某一变量的偏导与该变量不能同时非零,如  $f'(\lambda_i)\lambda_i=0$ 。注意所涉及变量要对应

约束规范: 边界的某些不规则性可能会使 KT 条件失效

充分性定理

# 集值映射下最优值的存在与连续性

考虑如下形式的最大化问题

$$M(a) = \max f(\mathbf{x}, a)$$

$$s.t. \ \mathbf{x} \in G(a)$$
(29)

### 最优值的存在性

如果约束集 G(a) 是非空且紧的,函数 f 是连续的,那么这个最大化问题存在一个解  $\mathbf{x}^*$  。注意  $R^n$  中的紧集也就是非空闭集,可能可以理解为类似于闭区间上连续函数必有最值

### 最优值的唯一性

如果函数 f 是严格凹的,且约束集是凸的,那么,若解存在,则它是唯一的

### 半连续性

对应(correspondence),即经济学家口中的多值函数。集值映射

### 最大值定理

令  $f(\mathbf{x},a)$  为有一紧值域的一个连续函数,假设约束集 G(a) 是 a 的一个非空、紧值、连续的对应,那么:

- 1. 最大值 M(a) 是一个连续函数
- 2. 最大化问题中的解  $\mathbf{x}(a)$  是一个上半连续的对应

如果对应  $\mathbf{x}(a)$  恰好是单值的,使得  $\mathbf{x}(a)$  为一个函数,那么它将会是一个连续函数

## 可积性条件

对于一般如下形式带边界条件的偏微分方程组

$$\frac{\partial f(\mathbf{p})}{\partial p_i} = g_i(f(\mathbf{p}, \mathbf{p})) \quad i = 1, \dots, n$$

$$f(\mathbf{q}) = 0$$
(30)

局部解存在的必要条件是交叉偏导数的对称性

$$\frac{dg_i}{dp_j} = \frac{dg_i}{\partial f} \frac{\partial f}{\partial p_j} + \frac{\partial g_i}{\partial p_j} = \frac{\partial^2 f(\mathbf{p})}{\partial p_i \partial p_j} \stackrel{\text{\tiny X is.}}{=} \frac{i,j}{\partial p_j \partial p_i} = \frac{dg_j}{\partial f} \frac{\partial f}{\partial p_i} + \frac{\partial g_j}{\partial p_i} = \frac{dg_j}{dp_i}$$
(31)

# 动态最优化

允许控制变量路径出现第一型间断点。也就是说允许状态变量路径出现尖点,只要分段可微

### 汉密尔顿函数 Hamiltonian Function

最大化典型问题(最小化就要另加负号),有自由终止状态(垂直终止线)

$$Max\ V = \int_0^T F(t,y,u)dt$$
 (32) 
$$s.t.\ \frac{dy}{dt} = f(t,y,u)$$
  $y(0) = A, y(T)$ 自由( $A, T$ 给定) 
$$u(t) \in \mathcal{U}$$
对于所有 $t \in [0,T]$ 

其中

- t, 时间变量
- y, 状态变量
- u,控制变量
- λ, 共态变量

汉密尔顿函数

$$H(t, y, u, \lambda) \equiv F(t, y, u) + \lambda(t)f(t, y, u)$$
(33)

最大值原理(一阶必要条件)

$$Max\ H(t,y,u,\lambda)$$
 对于所有 $t\in[0,T]$  
$$\frac{\partial H}{\partial \lambda} = \frac{dy}{dt} \ (y$$
的运动方程,照抄约束条件) 
$$\frac{\partial H}{\partial y} = -\frac{d\lambda}{dt} \ (\lambda$$
的运动方程,注意负号) 
$$\lambda(T) = 0 \ (横截条件)$$

注意第一个条件,是个比  $\frac{\partial H}{\partial u}=0$  更宽泛的要求,因为可能会出现边界解。严格点还能再用二阶导数确定是最大值而不是最小值

### 横截条件

• 垂直终止线:  $\lambda(T)=0$ , Ramsey model 常用就是这个

• 固定终止点:  $y(T)=y_T$ 

• 水平终止线:  $H_{t=T}=0$ 

• 终止曲线  $y_T = \phi(T)$ :  $[H - \lambda \phi']_{t=T} = 0$ 

• 截断垂直终止线:  $\lambda(T) \geq 0, y_T \geq y_{min}, (y_T - y_{min})\lambda(T) = 0$ 

• 截断水平终止线:  $H_{t=T} \geq 0, T \leq T_{max}, (T - T_{max})H_{t=T} = 0$ 

求解时整合上面各式,要求出**控制变量**的均衡值,或者至少是其路径;状态变量也可以对应求出

### 当前值汉密尔顿函数

被积函数增加了连续贴现因子 $e^{-\rho t}$ ,最优控制问题变为

$$Max \ V = \int_0^T G(t,y,u)e^{-\rho t}dt$$
 (35) 
$$s.t. \ \frac{dy}{dt} = f(t,y,u)$$
  $y(0) = A, y(T)$ 自由( $A, T$ 给定) 
$$u(t) \in \mathcal{U}$$
对于所有 $t \in [0,T]$ 

定义一个新的当前值拉格朗日乘子  $m=\lambda e^{\rho t}$  从而有当前值汉密尔顿函数

$$H_c \equiv He^{\rho t} = G(t, y, u) + mf(t, y, u) \tag{36}$$

### 当前值下的最大值原理

#### 无穷水平下的问题

横截条件仍然和有限的一样,只要取  $lim_{T o +\infty}$  即可(应该)

# 欧拉方程

# 离散形式

两期间先转移一个微分量,然后这一转移带来的两期边际效用增减必须相等

例:某一资产定价为  $P_t$ ,每一期(在购买/出售交易前)会产生孳息  $D_t$ 。利率外生为 r,效用贴现率为  $\rho$ ,瞬时消费效用为 CRRA 变体  $u(C_t)=lnC_t$ 。则考虑某一期减少消费去购买此资产,下一期出手,可以列出欧拉方程

$$dC \cdot dlnC_{t} = E_{t} \left[ \frac{1}{1+\rho} \frac{dC}{P_{t}} (D_{t+1} + P_{t+1}) \cdot dlnC_{t+1} \right]$$
(38)

如果不考虑购买资产, 只是两期间消费转移

$$dC \cdot dlnC_t = (1+r)\frac{1}{1+\rho} \cdot dlnC_{t+1}$$
(39)

### 动态最优化形式

### 反推最优化问题