# 第四章 賦範空間中的微分

## 第一節 微分與偏微分

在大一微積分中，我們有討論過在給定區間 $I \subseteq \mathbb{R}$ 時，函數 $f : I \to \mathbb{R}$ 微分的概念。這當中，最重要的是，我們能夠寫下 $f$ 在 $x \in I$ 附近的泰勒展開式：

$$f(x+h) = f(x) + hf'(x) + o(h),$$

其中 $h \mapsto hf'(x)$ 是函數 $f$ 在 $x$ 附近線性化所得到的結果。如果我們考慮的函數取值在高維度的歐氏空間 $\mathbb{R}^n$ 中，我們也能夠發展類似的理論。接下來，我們就要看怎麼把這套理論推廣到更一般的函數，定義域是賦範向量空間的開集，值域則是另一個賦範空間。

### 第一小節 微分

令 $(V, \|\cdot\|_V)$ 及 $(W, \|\cdot\|_W)$ 為兩個賦範向量空間。我們考慮開集 $A \subseteq V$ 還有 $f : A \to W$。

**定義 4.1.1 ：** 令 $a \in A$。如果存在 $\varphi \in \mathcal{L}_c(V, W)$ 使得

$$f(a+h) = f(a) + \varphi(h) + o(\|h\|_V), \quad 當 \quad h \to 0, \tag{4.1}$$

則我們說 $f$ 在 $a$ 是可微的[1]。如果映射 $\varphi$ 存在，那麼他是唯一的，我們把他稱作 $f$ 在 $a$ 的微分 (differential)，記作 $Df(a)$ 或 $\mathrm{d}f_a$。

**註解 4.1.2 ：** 由於 $A$ 是開集，且 $a$ 是個內點，對於夠靠近 $0$ 的 $h$，我們知道 $a+h$ 也會在 $A$ 裡面。因此，式 (4.1) 中的條件「$h \to 0$」有他存在的重要性，因為這個式子只有在當 $h$ 夠靠近 $0$ 時才有意義。

**定義 4.1.3 ：** 如果 $f$ 在所有的 $a \in A$ 皆可微，我們說 $f$ 在 $A$ 上是可微的，我們把

$$\begin{array}{rccc} Df: & A & \to & \mathcal{L}_c(V, W) \\ & a & \mapsto & \mathrm{d}f_a \end{array}$$

稱作 $f$ 的微分函數。如果 $Df$ 是連續的，則我們說 $f$ 是個 $\mathcal{C}^1$ 類的函數。

[1]也稱作 Fréchet 可微，在習題 4.10 中我們會討論到更一般的可微性，稱作 Gâteaux 可微。

**註解 4.1.4 ：**

(1) 如果 $V = \mathbb{R}$，這裡所討論的微分 (differential) 與在一維上的微分 (derivative) 的概念是相同的，換句話說，映射 $Df(a)$ 可以寫做 $Df(a)(h) = \mathrm{d}f_a(h) = f'(a)h$。所以我們可以直接記 $Df(a) = \mathrm{d}f_a = f'(a)$。

(2) 一般來說，微分 $\mathrm{d}f_a$ 的定義會取決於範數 $\|\cdot\|_V$ 還有 $\|\cdot\|_W$。然而，如果 $V$ 和 $W$ 是有限維度的向量空間，我們在定理 3.2.22 中看到，所有的範數皆是等價的，因此，微分 $\mathrm{d}f_a$ 的存在性以及他的取值，不取決於我們在空間上所賦予的範數。

(3) 要求微分 $\mathrm{d}f_a$ 是個連續映射是很重要的。在有限維度的空間中，所有線性映射皆是連續的（系理 3.2.24），因此在這樣的空間中，我們只需要檢查線性，就能自動得到連續性。

**範例 4.1.5 ：**

(1) 如果 $f \in \mathcal{L}_c(V, W)$，那麼透過關係式 $f(a+h) = f(a) + f(h)$，我們知道 $f$ 在 $V$ 上可微，且對於所有 $a \in V$，我們有 $\mathrm{d}f_a = f$。

(2) 考慮在 $\mathbb{R}^2$ 上的乘法：

$$\begin{array}{rccc} \psi: & \mathbb{R}^2 & \to & \mathbb{R} \\ & (x,y) & \mapsto & xy \end{array}.$$

那麼我們有

$$\psi(x+h_x, y+h_y) - \psi(x,y) = xh_y + h_x y + h_x h_y.$$

由於映射 $(h_x, h_y) \mapsto xh_y + yh_x$ 是線性的，且 $h_x h_y = o(\|(h_x, h_y)\|)$，我們推得 $\mathrm{d}\psi_{x,y}(h) = xh_y + yh_x$ 對於 $h = (h_x, h_y) \in \mathbb{R}^2$。

(3) 考慮在 $\mathcal{M}_n(\mathbb{R})$ 上的矩陣乘法：

$$\begin{array}{rccc} \psi: & \mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R}) & \to & \mathcal{M}_n(\mathbb{R}) \\ & (M,N) & \mapsto & MN \end{array}.$$

我們在向量空間 $\mathcal{M}_n(\mathbb{R})$ 上賦予定義在註解 3.2.16 中的範數 $|||\cdot|||$。固定 $M, N \in \mathcal{M}_n(\mathbb{R})$。那麼，對於 $H, K \in \mathcal{M}_n(\mathbb{R})$，我們有

$$\psi(M+H, N+K) - \psi(M,N) = MK + HN + HK.$$

映射 $(H,K) \mapsto MK + HN$ 是線性的，且 $|||HK||| \leqslant |||H||| |||K||| \leqslant \|(H,K)\|^2$，因此我們得到 $\mathrm{d}\psi_{M,N}(H,K) = MK + HN$。

**範例 4.1.6 ：** 令 $V$ 為賦範向量空間，且

$$\mathcal{GL}_c(V) = \{u \in \mathcal{L}(V,V) : u \text{ 以及 } u^{-1} \text{ 皆連續}\}.$$

定義函數 $\mathrm{Inv} : \mathcal{GL}_c(V) \to \mathcal{GL}_c(V), u \mapsto u^{-1}$。對於 $h \in \mathcal{GL}_c(V)$ 滿足 $|||h||| < 1$，我們知道 $\mathrm{id} + h$ 是可逆的，且反函數寫做

$$(\mathrm{id} + h)^{-1} = \mathrm{id} - h + \sum_{n \geqslant 2} (-1)^n h^n.$$

我們有

$$\left|\left|\left|\sum_{n \geqslant 2} (-1)^n h^n\right|\right|\right| \leqslant \sum_{n \geqslant 2} |||h|||^n = \frac{|||h|||^2}{1 - |||h|||}.$$

因此，當 $h \to 0$，我們得到

$$(\mathrm{id} + h)^{-1} = \mathrm{id} - h + o(|||h|||).$$

這代表著 Inv 在 id 是可微的，且他的微分寫做 $\mathrm{d}\,\mathrm{Inv}_{\mathrm{id}} : h \mapsto -h$。

**命題 4.1.7 ：** 如果 $f$ 在 $a \in A$ 可微，那麼 $f$ 也會在 $a$ 連續。

**證明：**假設 $f$ 在 $a \in A$ 可微，那麼我們能找到連續線性函數 $\varphi : V \to W$ 以及 $r > 0$ 使得

$$\forall h \in B_V(0,r), \quad f(a+h) = f(a) + \varphi(h) + \|h\|_V\, \varepsilon(h),$$

其中 $\lim_{h \to 0} \varepsilon(h) = 0$。固定 $\delta > 0$ 以及 $0 < r' \leqslant r$ 使得 $\|\varepsilon(h)\|_W < \delta$ 對於 $h \in B_V(0,r')$。那麼，我們有

$$\forall h \in B_V(0,r'), \quad \|f(a+h) - f(a)\|_W \leqslant \|\varphi(h)\|_W + \|h\|_V\, \delta \leqslant (M + \delta)\, \|h\|_V\,,$$

其中 $M = |||\varphi|||$。這給我們 $f$ 在 $a$ 的連續性。 □

**命題 4.1.8 ：** 令 $V, W$ 為兩個賦範向量空間，$A \subseteq V$ 為開子集合，且 $f, g : A \to W$ 為兩個在 $a \in A$ 可微的函數。那麼，我們有：

(1) $f + g$ 在 $a$ 可微，且 $\mathrm{d}(f+g)_a = \mathrm{d}f_a + \mathrm{d}g_a$；

(2) 對於所有 $\lambda \in \mathbb{K}$，函數 $\lambda f$ 在 $a$ 可微，且 $\mathrm{d}(\lambda f)_a = \lambda\, \mathrm{d}f_a$。

**證明：** 請使用定義 4.1.1 中的定義，自己完成此命題的證明。 □

**命題 4.1.9【鏈鎖律】：** 令 $V, W, X$ 為在 $\mathbb{K}$ 上的賦範向量空間，$A \subseteq V$ 以及 $B \subseteq W$ 為兩個開子集。考慮兩個函數 $f : A \subseteq V \to W$ 以及 $g : B \subseteq W \to X$ 滿足 $f(A) \subseteq B$。假設 $f$ 在 $a \in A$ 可微，且 $g$ 在 $f(a)$ 可微，那麼 $g \circ f : A \subseteq V \to X$ 在 $a$ 可微，且我們有

$$\mathrm{d}(g \circ f)_a = \mathrm{d}g_{f(a)} \circ \mathrm{d}f_a. \tag{4.2}$$

**註解 4.1.10：** 如果 $V = W = X = \mathbb{R}$，那麼式 (4.2) 可以重新寫做 $(g \circ f)'(a) = g'(f(a)) \cdot f'(a)$，這是我們在大一微積分所看過的鏈鎖律。

**證明：** 根據 $f$ 在 $a$ 的可微性，我們有

$$f(a+h) = f(a) + \mathrm{d}f_a(h) + o(\|h\|_V), \quad \text{當 } h \to 0.$$

當我們把他與 $g$ 合成，再使用 $g$ 在 $b = f(a)$ 的可微性，我們得到

$$\begin{aligned}(g \circ f)(a+h) &= g(\underbrace{f(a)}_{b} + \underbrace{\mathrm{d}f_a(h) + o(\|h\|_V)}_{h'}) \\ &= g(f(a)) + \mathrm{d}g_b(h') + o(\|h'\|_V).\end{aligned}$$

由於 $\mathrm{d}f_a \in \mathcal{L}_c(V, W)$，根據定理 3.2.12，我們知道 $h' = O(\|h\|_V)$。相似地，根據 $\mathrm{d}g_b \in \mathcal{L}_c(W, X)$，我們會有

$$\mathrm{d}g_b(h') = \mathrm{d}g_b \circ \mathrm{d}f_a(h) + \mathrm{d}g_b(o(\|h\|_V)) = \mathrm{d}g_b \circ \mathrm{d}f_a(h) + o(\|h\|_V),$$

且 $\mathrm{d}g_b \circ \mathrm{d}f_a$ 是個連續線性映射，因為他是連續線性映射的合成函數。因此，

$$(g \circ f)(a+h) = (g \circ f)(a) + \mathrm{d}g_b \circ \mathrm{d}f_a(h) + o(\|h\|_V), \quad \text{當 } h \to 0.$$

這讓我們推得 $\mathrm{d}(g \circ f)_a = \mathrm{d}g_b \circ \mathrm{d}f_a$。 □

**系理 4.1.11：** 令 $f, g : A \subseteq V \to \mathbb{R}$ 在 $a \in A$ 可微，那麼他們的積 $fg$ 也會在 $a$ 可微，且我們有

$$\mathrm{d}(fg)_a = g(a) \cdot \mathrm{d}f_a + f(a) \cdot \mathrm{d}g_a.$$

**證明：**這是可以透過使用命題 4.1.9 而直接得到的結果。我們考慮函數

$$\begin{array}{rccc}\varphi: & A & \to & \mathbb{R}^2 \\ & x & \mapsto & (f(x),g(x))\end{array}, \quad \text{且} \quad \begin{array}{rccc}\psi: & \mathbb{R}^2 & \to & \mathbb{R} \\ & (x,y) & \mapsto & xy\end{array}.$$

那麼，乘積 $fg$ 會是合成函數 $x \mapsto (\psi \circ \varphi)(x)$，且我們有

$$\begin{aligned} \mathrm{d}\varphi_x(h) &= (\mathrm{d}f_x(h), \mathrm{d}g_x(h)) \\ \mathrm{d}\psi_{x,y}(h_x,h_y) &= h_x y + h_y x. \end{aligned}$$

因此，藉由合成，當 $h \to 0$ 時，我們得到

$$\mathrm{d}(fg)_a(h) = \mathrm{d}\psi_{\varphi(a)} \circ \mathrm{d}\varphi_a(h) = g(a)\,\mathrm{d}f_a(h) + f(a)\,\mathrm{d}g_a(h)$$

□

## 第二小節 均值定理

我們回顧在大一微積分所看過，對於連續可微函數 $f : I \to \mathbb{R}$，其中 $I \subseteq \mathbb{R}$ 是個開區間，我們有下面的均值定理。對於 $a, b \in I$ 滿足 $a < b$，存在 $c \in (a,b)$ 使得

$$f(b) - f(a) = f'(c)(b-a). \tag{4.3}$$

此外，如果我們知道 $\sup_{t\in[a,b]} |f'(c)| \leqslant M$，那麼 $|f(b)-f(a)| \leqslant M(b-a)$，我們稱之為均值不等式。在下面，我們會把均值定理還有均值不等式，推廣到更一般的情況，也就是說，我們會對定義在賦範向量空間開集，取值是在另一個賦範向量空間的函數做討論。

**引理 4.1.12 ：** 令 $a < b$ 為實數，且 $W$ 是個賦範向量空間。令 $f : [a,b] \to W$ 還有 $g : [a,b] \to \mathbb{R}$ 為兩個在 $[a,b]$ 上的連續函數，且在 $(a,b)$ 上可微。如果對於所有 $t \in (a,b)$，我們有 $\|f'(t)\|_W \leqslant g'(t)$，那麼 $\|f(b)-f(a)\|_W \leqslant g(b)-g(a)$。

**證明：**首先，讓我們假設 $\|f'(t)\|_W < g'(t)$ 對於所有 $t \in (a,b)$。這代表著

$$\begin{aligned} & \forall t \in (a,b), \quad \lim_{\substack{x\to t \\ x>t}} \left\| \frac{f(x)-f(t)}{x-t} \right\|_W - \frac{g(x)-g(t)}{x-t} < 0 \\ \Rightarrow\quad & \forall t \in (a,b), \exists y > t, \forall x \in (t,y), \quad \left\| \frac{f(x)-f(t)}{x-t} \right\|_W < \frac{g(x)-g(t)}{x-t} \\ \Rightarrow\quad & \forall t \in (a,b), \exists y > t, \forall x \in [t,y], \quad \|f(x)-f(t)\|_W \leqslant g(x)-g(t). \end{aligned} \tag{4.4}$$

令 $[\alpha,\beta] \subseteq (a,b)$，我們想要證明

$$\|f(\beta)-f(\alpha)\|_W \leqslant g(\beta)-g(\alpha). \tag{4.5}$$

令

$$\Gamma = \{\theta \in (\alpha, \beta] : \forall x \in [\alpha, \theta], \|f(x) - f(\alpha)\|_W \leqslant g(x) - g(\alpha)\}.$$

根據式 (4.4)，我們知道 $\Gamma$ 是非空的。令 $\gamma = \sup \Gamma$，如果我們能證明 $\gamma = \beta$，我們就能得到式 (4.5)。

我們使用反證法。假設 $\gamma < \beta$。由於 $f$ 和 $g$ 皆是連續的，我們也會有

$$\|f(\gamma) - f(\alpha)\|_W \leqslant g(\gamma) - g(\alpha). \tag{4.6}$$

但從式 (4.4)，我們知道

$$\exists \delta \in (\gamma, \beta], \forall x \in [\gamma, \delta], \quad \|f(x) - f(\gamma)\|_W \leqslant g(x) - g(\gamma). \tag{4.7}$$

接著，把式 (4.6) 還有式 (4.7) 放在一起，並使用三角不等式，我們知道會存在 $\delta \in (\gamma, \beta]$ 使得

$$\forall x \in [\gamma, \delta], \quad \|f(x) - f(\alpha)\|_W \leqslant g(x) - g(\alpha).$$

這讓我們得出 $\delta \in \Gamma$，這是不可能的，因為我們前面假設 $\delta > \gamma = \sup \Gamma$。所以我們得知式 (4.5) 為真。接著，在式 (4.5) 中，我們可以取 $\alpha \to a$ 還有 $\beta \to b$，透過 $f$ 和 $g$ 的連續性，我們也會有 $\|f(b) - f(a)\|_W \leqslant g(b) - g(a)$。

最後我們總結，我們所需要處理的情況是原本的假設 $\|f'(t)\|_W \leqslant g'(t)$ 對於所有 $t \in (a, b)$。固定 $\varepsilon > 0$，我們可以考慮函數 $g_\varepsilon(t) = g(t) + \varepsilon t$ 其中 $t \in [a, b]$。那麼對於 $t \in (a, b)$，我們有 $\|f'(t)\|_W < g'_\varepsilon(t)$。上半部的證明讓我們推得 $\|f(b) - f(a)\|_W \leqslant g_\varepsilon(b) - g_\varepsilon(a)$。取 $\varepsilon \to 0$，我們得到我們想要證明的結果。 □

**定理 4.1.13 【均值不等式】：** 令 $V$ 及 $W$ 為兩個賦範向量空間，且 $A \subseteq V$ 是個開子集。令 $f : A \subseteq V \to W$ 為函數。考慮 $a, b \in A$ 使得線段 $[a, b] \subseteq A$。假設

(a) $f$ 在 $[a, b]$ 上連續；

(b) $f$ 在 $(a, b)$ 上可微；

(c) 存在 $M > 0$ 使得 $\|\mathrm{d}f_c\| \leqslant M$ 對於 $c \in (a, b)$。

那麼，我們有

$$\|f(b) - f(a)\|_W \leqslant M \|b - a\|_V . \tag{4.8}$$

**證明：**考慮函數 $g : [0, 1] \to W$ 定義為 $g(t) = f(a + t(b - a))$ 對於 $t \in [0, 1]$。那麼，$g$ 在 $[0, 1]$ 上連續且在 $(0, 1)$ 上可微，他的微分寫做

$$g'(t) = \mathrm{d}f_{a+t(b-a)}(b - a), \quad \forall t \in (a, b).$$

因此，對於 $t \in (0,1)$，我們有 $\|g'(t)\|_W \leqslant M \|b-a\|_V$。根據引理 4.1.12 ，我們得到想要的結果。 □

**註解 4.1.14 ：** 我們注意到，在一般的賦範向量空間（維度大於等於 2），我們能得到的最好結果就只是個不等式而已，即使是在定理 4.1.13 中的條件 (c) 中，微分的算子範數恆等於 $M$ 時也是一樣。例如，我們可以考慮下面這個映射

$$\begin{array}{rccc} f: & \mathbb{R} & \to & \mathbb{R}^2 \\ & t & \mapsto & (\cos t, \sin t). \end{array}$$

我們不難檢查，對於所有 $t \in \mathbb{R}$，我們有 $\mathrm{d}f_t = (-\sin t, \cos t)$ 且會滿足 $\|\!|\mathrm{d}f_t|\!\| = 1$。然而，我們有 $\|f(0) - f(2\pi)\| = 0 \neq 2\pi \cdot 1$。

**定理 4.1.15 【均值定理】：** 令 $V$ 為賦範向量空間，$W = \mathbb{R}^n$ 為歐氏空間，且 $A \subseteq V$ 是個開子集合。考慮在 $A$ 上可微的函數 $f : A \subseteq V \to \mathbb{R}^n$。令 $a, b \in A$ 使得 $[a,b] \subseteq A$。那麼，對於任意向量 $v \in \mathbb{R}^n$，存在 $c \in (a,b)$ 使得

$$v \cdot [f(b) - f(a)] = v \cdot \mathrm{d}f_c(b-a). \tag{4.9}$$

**證明：** 令 $h = b - a$。由於 $A$ 是開集，且 $[a, a+h] \subseteq A$，存在 $\delta > 0$ 使得 $a + th \in A$ 對於 $t \in (-\delta, 1+\delta)$。固定向量 $v \in \mathbb{R}^n$ 並令 $g : (-\delta, 1+\delta) \to \mathbb{R}$ 定義做

$$g(t) = v \cdot f(a+th), \quad \forall t \in (-\delta, 1+\delta).$$

那麼，$f$ 在 $(-\delta, 1+\delta)$ 上是可微的，且他的微分寫做

$$g'(t) = v \cdot \mathrm{d}f_{a+th}(h).$$

根據一維的均值定理（式 (4.3)），我們得到

$$g(1) - g(0) = g'(t), \quad \text{對於某個 } t \in (0,1),$$

這正好就是式 (4.9)。 □

## 第三小節　方向導數

**定義 4.1.16 ：** 令 $a \in A$ 以及向量 $u \in V$。如果下面極限存在

$$f'_u(a) = \lim_{h \to 0} \frac{f(a+hu) - f(a)}{h}, \tag{4.10}$$

則我們說 $f$ 在 $a$ 沿著 $u$ 的方向導數存在，記作 $f'_u(a)$。

**命題 4.1.17 ：** 如果 $f$ 在 $a$, 可微，那麼他在 $a$ 沿著任意方向 $u \in V$ 的微分定義良好，且我們有 $f'_u(a) = \mathrm{d}f_a(u) = Df(a)(u)$。

**註解 4.1.18 ：** 我們注意到，如果 $f$ 在 $a$ 對於任意方向的方向導數皆存在，那不一定代表著 $f$ 在 $a$ 是可微的；實際上，他可以甚至在 $a$ 不連續。我們可以考慮 $f : \mathbb{R}^2 \to \mathbb{R}$ 定義做

$$f(x,y) = \begin{cases} \frac{y^2}{x}, & 若\ x \neq 0, \\ y, & 若\ x = 0. \end{cases}$$

那麼，$f$ 在 $(0,0)$ 不連續，因為我們有

$$\lim_{x\to 0} f(x, \sqrt{x}) = 1 \neq 0 = f(0,0).$$

然而，對於任意 $u = (a,b) \in \mathbb{R}^2$，$f$ 在 $(0,0)$ 沿著方向 $u$ 的方向導數存在：

$$f'_u(0,0) = \lim_{h\to 0} \frac{f(h(a,b)) - f(0,0)}{h} = \begin{cases} \frac{b^2}{a}, & 若\ a \neq 0, \\ b, & 若\ a = 0. \end{cases}$$

接下來，我們取 $V = \mathbb{R}^n$ 為 $n$ 維度的歐氏空間，他的標準基底記作 $(e_1, \ldots, e_n)$。令 $A$ 為 $V$ 的開子集合，以及 $f : A \to W$。

**定義 4.1.19 ：** 對於 $1 \leqslant i \leqslant n$，如果 $f$ 在 $a$ 沿著方向 $e_i$ 的方向導數存在，則我們說他在 $a$ 對於第 $i$ 個座標的偏微分存在，且定義

$$\frac{\partial f}{\partial x_i}(a) = f'_{e_i}(a) \tag{4.11}$$

**註解 4.1.20 ：**

(1) 與註解 4.1.18 中所提到的類似，我們能夠找到 $f$ 使得他在 $a$ 的所有偏微分存在，但 $f$ 在 $a$ 不是可微，甚至是不連續的。

(2) 如果 $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ 在 $a \in A$ 可微，那麼他所有在 $a$ 的偏微分皆存在，且我們有

$$Df(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\,\mathrm{d}x_i, \qquad \overrightarrow{\mathrm{grad}}_a\, f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) e_i,$$

其中 $(\mathrm{d}x_i = e_i^*)_{1\leqslant i\leqslant n}$ 是 $(\mathbb{R}^n)^* = \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ 中，相對應到 $\mathbb{R}^n$ 中標準基底 $(e_i)_{1\leqslant i\leqslant n}$ 的對偶基底；換句話說

$$\mathrm{d}x_i(e_j) = e_i^*(e_j) = \delta_{i,j}, \quad \forall 1 \leqslant i,j \leqslant n.$$

這會給我們：

$$Df(a)(h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) h_i = \left(\overrightarrow{\mathrm{grad}}_a f\right) \cdot h. \tag{4.12}$$

**定理 4.1.21 ：** 令 $f : A \subseteq \mathbb{R}^n \to W$。假設

(a) 在 $A$ 上 $f$ 所有的偏微分存在；

(b) 所有偏微分在 $a$ 連續。

那麼，$f$ 在 $a$ 可微，且我們有

$$Df(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\, \mathrm{d}x_i. \tag{4.13}$$

**註解 4.1.22 ：** 我們重新提醒一次，這裡 $Df(a)$ 是個從 $\mathbb{R}^n$ 映射至 $W$ 的線性函數。對於每個 $1 \leqslant i \leqslant n$，偏微分 $\frac{\partial f}{\partial x_i}(a)$ 是個在 $W$ 中的向量，$\mathrm{d}x_i$ 是個在 $\mathbb{R}^n$ 上的線性泛函，也就是說，是個由 $\mathbb{R}^n$ 映射至 $\mathbb{R}$ 的線性（連續）函數。如果我們把式 (4.13) 取值在 $u \in \mathbb{R}^n$，左手邊會給我們 $Df(a)(u) \in W$，右手邊中，每一項會給我們純量 $\mathrm{d}x_i(u) = u_i \in \mathbb{R}$，乘上向量 $\frac{\partial f}{\partial x_i}(a) \in W$。

**證明：**我們在 $\mathbb{R}^n$ 上賦予範數 $\|x\| = \sum_{i=1}^{n} |x_i|$。令

$$\begin{aligned} g : \quad A &\to W \\ x &\mapsto f(x) - \sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i}(a). \end{aligned}$$

我們想要證明，當 $x \to a$ 時，我們有 $g(x) - g(a) = o(\|x - a\|)$。

令 $\varepsilon > 0$。(b) 中對於連續性的假設告訴我們存在 $r > 0$ 使得對於所有 $1 \leqslant i \leqslant n$，我們有

$$\forall x \in A \cap B(a,r), \quad \left\| \frac{\partial g}{\partial x_i}(x) \right\|_W = \left\| \frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(a) \right\|_W < \varepsilon. \tag{4.14}$$

由於 $A$ 是個開集，我們可以選擇更小的 $r > 0$，以便也假設 $B(a,r) \subseteq A$。

對於 $x \in B(a,r)$，我們考慮下面的點

$$\begin{aligned} y_0 &= (a_1, \ldots, a_n) = a, \\ y_k &= (x_1, \ldots, x_k, a_{k+1}, \ldots, a_n), \quad \forall k = 1, \ldots, n. \end{aligned}$$

我們注意到 $y_0 = a$, $y_n = x$，且中間的點 $y_k$ 是由透過把 $a$ 中的座標，一一置換為 $x$ 的座標所得

到的。對於 $1 \leqslant k \leqslant n$，定義

$$\begin{array}{rccc} g_k: & [a_k, x_k] & \to & W \\ & t & \mapsto & g(x_1, \ldots, x_{k-1}, t, a_{k+1}, \ldots, a_n). \end{array}$$

函數 $g_k$ 的導數寫做

$$g_k'(t) = \frac{\partial g}{\partial x_k}(x_1, \ldots, x_{k-1}, t, a_{k+1}, \ldots, a_n),$$

再根據式 (4.14)，我們得知在 $[a_k, x_k]$ 上，我們有 $\|g_k'(t)\|_W < \varepsilon$。因此，從引理 4.1.12 ，我們得到

$$\|g_k(x_k) - g_k(a_k)\|_W \leqslant \varepsilon |x_k - a_k|.$$

由於 $g_k(a_k) = g(y_{k-1})$ 且 $g_k(x_k) = g(y_k)$，我們有

$$\begin{aligned} \|g(x) - g(a)\|_W &= \left\| \sum_{k=1}^{n} [g(y_k) - g(y_{k-1})] \right\|_W \leqslant \sum_{k=1}^{n} \|g(y_k) - g(y_{k-1})\|_W \\ &\leqslant \varepsilon \sum_{k=1}^{n} |x_k - a_k| = \varepsilon \|x - a\| . \end{aligned}$$

因此，我們得到

$$\forall x \in B(a, r), \quad \|g(x) - g(a)\|_W \leqslant \varepsilon \|x - a\| .$$

這個式子也可以等價寫做 $g(x) - g(a) = o(\|x - a\|)$。 $\square$

**註解 4.1.23 ：** 注意到定理 4.1.21 的逆命題是錯誤的。我們可以找到可微函數，但偏微分是不連續的。例如，考慮下面這個經典的範例 $f : \mathbb{R} \to \mathbb{R}$，定義做

$$f(x) = \begin{cases} x^2 \sin(\frac{1}{x}), & \text{若 } x \neq 0, \\ 0, & \text{若 } x = 0. \end{cases}$$

我們可以計算 $f$ 在 $0$ 的微分：

$$f'(0) = \lim_{h \to 0} \frac{f(h) - f(0)}{h} = \lim_{h \to 0} h \sin(\tfrac{1}{h}) = 0.$$

然而 $f$ 在 $x \neq 0$ 的微分寫做：

$$f'(x) = 2x \sin(\tfrac{1}{x}) - \cos(\tfrac{1}{x}).$$

顯然地，微分函數 $f'$ 在 $0$ 不連續。

## 第四小節 Jacobi 矩陣

我們考慮特別的情況，把我們的賦範向量空間取做歐氏空間，也就是 $V=\mathbb{R}^n$ 還有 $W=\mathbb{R}^m$，對於特定的 $n,m\geqslant 1$。令 $(v_1,\ldots,v_n)$ 為 $V=\mathbb{R}^n$ 的標準基底，$(w_1,\ldots,w_m)$ 為 $W=\mathbb{R}^m$ 的標準基底。令 $A\subseteq\mathbb{R}^n$ 為開子集合，以及 $f:A\to\mathbb{R}^m$ 為在 $a\in A$ 可微的函數。由於 $\mathrm{d}f_a\in\mathcal{L}_c(\mathbb{R}^n,\mathbb{R}^m)$，如果把他寫在標準基底中，他也可以用一個 $m\times n$ 實係數的矩陣來表示，換句話說，這個矩陣的係數分別是：

$$\mathrm{d}f_a(v_j)\cdot w_i,\quad 1\leqslant i\leqslant m, 1\leqslant j\leqslant n.$$

**定義 4.1.24：** $f$ 在 $a$ 的 *Jacobi 矩陣*是矩陣 $J_f(a)\in\mathcal{M}_{m,n}(\mathbb{R})$，寫做

$$J_f(a)=\left[\frac{\partial f_i}{\partial x_j}(a)\right]_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}},$$

其中對於 $1\leqslant i\leqslant m$，我們記 $f_i=\mathrm{Proj}_i\circ f$，所以會有 $f=\sum_{i=1}^m f_iw_i$。當 $m=n$ 時，Jacobi 矩陣是個方形矩陣，我們把他的行列式 $\det(J_f(a))$ 稱為 *Jacobi 行列式*，或是 *Jacobian*。

**註解 4.1.25：** 我們注意到，Jacobi 矩陣 $J_f(a)$ 的 $i$ 列，會是函數 $f_i$ 的梯度，也就是說

$$\overrightarrow{\mathrm{grad}}_a f_i=\sum_{j=1}^n\frac{\partial f_i}{\partial x_j}(a)v_j,\quad\text{或}\quad\left(\overrightarrow{\mathrm{grad}}_a f_i\right)_{v_1,\ldots,v_n}=\left(\frac{\partial f_i}{\partial x_j}(a)\right)_{1\leqslant j\leqslant n}.$$

使用式 (4.12)，我們也可以把 $f$ 在 $a$ 的微分重新改寫，對於任意 $h\in\mathbb{R}^n$，我們會有

$$Df(a)(h)=\sum_{i=1}^m Df_i(a)(h)w_i=\sum_{i=1}^m\left[\left(\overrightarrow{\mathrm{grad}}_a f_i\right)\cdot h\right]w_i.$$

而這個剛好就是矩陣 $J_f(a)$ 與 $h$ 之間的乘積，其中向量 $h$ 可以標準基底 $(v_1,\ldots,v_n)$ 中，寫成 $n\times 1$ 的行矩陣，而最後得到的結果會是個 $m\times 1$ 的矩陣，會是 $Df(a)(h)$ 寫在 $\mathbb{R}^m$ 的標準基底 $(w_1,\ldots,w_m)$ 中所得到的。

**命題 4.1.26【合成函數與 Jacobi 矩陣】：** 令 $m,n,k\geqslant 1$ 以及 $A\subseteq\mathbb{R}^m$ 和 $B\subseteq\mathbb{R}^n$ 為兩個開子集合。令 $f:A\to\mathbb{R}^n$ 以及 $g:\mathbb{R}^n\to\mathbb{R}^k$ 使得 $f(A)\subseteq B$。假設 $f$ 在 $a$ 可微，且 $g$ 在 $f(a)$ 可微。對於 $1\leqslant i\leqslant n$，我們把 $f_i=\mathrm{Proj}_i\circ f$ 記作函數 $f$ 的第 $i$ 個座標。那麼，函數 $h=g\circ f:A\to\mathbb{R}^k$ 在 $a$ 可微，且他的 Jacobi 矩陣寫做

$$J_h(a)=J_g(f(a))\cdot J_f(a).$$

換句話說，對於所有 $1 \leqslant j \leqslant m$，我們也能夠寫

$$\frac{\partial h}{\partial x_j}(a) = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_j}(a) \frac{\partial g}{\partial y_i}(f(a)).$$

**證明：** 這是可以透過命題 4.1.9 得到的直接結果，我們需要把裡面的關係式重新表示成定義 4.1.24 中 Jacobi 矩陣即可。 □

**範例 4.1.27 ：** 令 $f : \mathbb{R}^2 \to \mathbb{R}$ 為 $\mathcal{C}^1$ 類的函數。考慮下面映射

$$\begin{array}{rccc} \varphi : & \mathbb{R}_{>0} \times \mathbb{R} & \to & \mathbb{R}^2 \\ & (r, \theta) & \mapsto & (r\cos\theta, r\sin\theta). \end{array}$$

這樣的情況下，合成函數 $F = f \circ \varphi$ 是個 $\mathcal{C}^1$ 類函數，且可以被視為函數 $f$ 在極座標中的表示式。我們有

$$\begin{aligned} & J_F(r,\theta) = J_f(r\cos\theta, r\sin\theta) J_\varphi(r,\theta) \\ \Leftrightarrow \quad & \begin{pmatrix} \frac{\partial F}{\partial r} & \frac{\partial F}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} \frac{\partial \varphi_1}{\partial r} & \frac{\partial \varphi_1}{\partial \theta} \\ \frac{\partial \varphi_2}{\partial r} & \frac{\partial \varphi_2}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}. \end{aligned}$$

換句話說，

$$\frac{\partial f}{\partial x} = \cos\theta \frac{\partial F}{\partial r} - \frac{\sin\theta}{r}\frac{\partial F}{\partial \theta} \quad \text{且} \quad \frac{\partial f}{\partial y} = \sin\theta \frac{\partial F}{\partial r} + \frac{\cos\theta}{r}\frac{\partial F}{\partial \theta}.$$

## 第二節　高階導數

在此小節中，我們會專注在有限維度向量空間的情況。然而，稍後在第 4.2.3 小節中，我們會提到高階微分在一般賦範向量空間中的推廣。

### 第一小節　Schwarz 定理

令 $A$ 為 $\mathbb{R}^n$ 的開子集合，且 $f : A \to W$ 為函數。令 $p \geqslant 1$ 為整數，且 $1 \leqslant i_1, \ldots, i_p \leqslant n$。在存在性的假設之下，我們可以透過遞迴方式來定義 $p$ 階導數：

$$\frac{\partial^p f}{\partial x_{i_p} \ldots \partial x_{i_1}} = \frac{\partial}{\partial x_{i_p}} \Big( \frac{\partial^{p-1} f}{\partial x_{i_{p-1}} \ldots \partial x_{i_1}} \Big).$$

如果 $f$ 的所有一直到 $p$ 階的導數皆存在且在 $A$ 上連續，則我們說 $f$ 是 $\mathcal{C}^p$ 類的。

**定理 4.2.1** 【Schwarz 定理】： 令 $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ 為函數，其中 $A \subseteq \mathbb{R}^2$ 是個開子集合。假設偏微分

$$\frac{\partial^2 f}{\partial x \partial y} \quad \text{且} \quad \frac{\partial^2 f}{\partial y \partial x}$$

在 $A$ 上存在，且在 $a \in A$ 連續。那麼，我們有

$$\frac{\partial^2 f}{\partial x \partial y}(a) = \frac{\partial^2 f}{\partial y \partial x}(a). \tag{4.15}$$

**註解 4.2.2** ： 從上述定理我們得知，在存在性和連續性的假設之下，偏微分的順序並不重要。

**範例 4.2.3** ： 這是由 Peano 提出的例子。考慮函數 $f : \mathbb{R}^2 \to \mathbb{R}$，定義做

$$f(x,y) = \begin{cases} xy\frac{x^2-y^2}{x^2+y^2} & \text{若 } (x,y) \neq (0,0), \\ 0 & \text{若 } (x,y) = (0,0). \end{cases}$$

一方面，我們有

$$\frac{\partial f}{\partial x}(x,y) = \begin{cases} \frac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2} & \text{若 } (x,y) \neq (0,0), \\ 0 & \text{若 } (x,y) = (0,0). \end{cases}$$

所以對於 $y \in \mathbb{R}$，我們得到 $\frac{\partial f}{\partial x}(0,y) = -y$，並且會給我們

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = -1.$$

另一方面，我們有

$$\frac{\partial f}{\partial y}(x,y) = \begin{cases} \frac{x(x^4-4x^2y^2-y^4)}{(x^2+y^2)^2} & \text{若 } (x,y) \neq (0,0), \\ 0 & \text{若 } (x,y) = (0,0). \end{cases}$$

所以對於 $x \in \mathbb{R}$，我們得到 $\frac{\partial f}{\partial y}(x,0) = x$，並且會給我們

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = 1.$$

實際上，我們不難檢查，他們的二階導數不連續。對於 $(x,y) \neq (0,0)$，我們有

$$\frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{x^6 + 9x^4y^2 - 9x^2y^4 - y^6}{(x^2+y^2)^3},$$

這會給我們

$$\lim_{x \to 0} \frac{\partial^2 f}{\partial y \partial x}(x,0) = 1, \quad \text{以及} \quad \lim_{y \to 0} \frac{\partial^2 f}{\partial y \partial x}(0,y) = -1.$$

如果不想做計算，你也可以使用反對稱性的關係，來檢查不連續性。

**證明：**不失一般性，我們可以假設 $a=(0,0)\in A$。令 $h,k>0$ 使得 $[0,h]\times[0,k]\subseteq A$ 且

$$\delta(h,k)=f(h,k)-f(h,0)-f(0,k)+f(0,0).$$

考慮函數 $\varphi$ 定義做

$$\begin{aligned}\varphi:\ [0,h] &\to \mathbb{R}\\ x &\mapsto f(x,k)-f(x,0).\end{aligned}$$

那麼 $\delta(h,k)=\varphi(h)-\varphi(0)$。由於 $\varphi$ 在 $[0,h]$ 上連續，且在 $(0,h)$ 上可微，根據 $\mathbb{R}$ 上的均值不等式（式 (4.3)），會存在 $t_1\in(0,1)$ 使得

$$\delta(h,k)=h\varphi'(t_1h)=h\Big[\frac{\partial f}{\partial x}(t_1h,k)-\frac{\partial f}{\partial x}(t_1h,0)\Big].$$

函數 $y\mapsto\frac{\partial f}{\partial x}(t_1h,y)$ 在 $[0,1]$ 上連續，且在 $(0,1)$ 上可微，再次使用均值不等式，我們得知存在 $t_2\in(0,1)$ 使得

$$\delta(h,k)=hk\frac{\partial^2 f}{\partial y\partial x}(t_1h,t_2k). \tag{4.16}$$

如果我們考慮函數 $\psi$ 定義做

$$\begin{aligned}\psi:\ [0,k] &\to \mathbb{R}\\ y &\mapsto f(h,y)-f(0,y)\end{aligned}$$

並且使用與上面相同的步驟，我們能找到 $t_3,t_4\in(0,1)$ 使得

$$\delta(h,k)=hk\frac{\partial^2 f}{\partial x\partial y}(t_3h,t_4k). \tag{4.17}$$

把式 (4.16) 還有式 (4.17) 放在一起，並且取 $h,k\to 0$，根據偏微分在 $(0,0)$ 的連續性，我們知道他們在 $(0,0)$ 是相等的。 □

**系理 4.2.4 ：** 令 $f:A\subseteq\mathbb{R}^n\to\mathbb{R}^m$ 為函數，其中 $A\subseteq\mathbb{R}^n$ 是個開子集合。假設 $f$ 是 $\mathcal{C}^p$ 類的，那麼所有一直到 $p$ 階的導數皆不取決於微分的順序。因此，我們可以把這些偏微分寫成下列形式：

$$\frac{\partial^k f}{\partial x_1^{i_1}\dots\partial x_n^{i_n}}\quad\text{其中}\quad i_1+\dots+i_n=k\leqslant p.$$

## 第二小節 Hessian 矩陣

**定義 4.2.5 ：** 令 $f: A \subseteq \mathbb{R}^n \to \mathbb{R}$ 為函數，其中 $A \subseteq \mathbb{R}^n$ 是個開子集合。假設 $f$ 在 $a \in A$ 所有的二階導數皆存在。那麼，$f$ 在 $a$ 的 *Hessian* 矩陣可以定義做

$$H_f(a) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right]_{1 \leqslant i,j \leqslant n}. \tag{4.18}$$

如果二階導數在 $a$ 都是連續的，根據 Schwarz 定理（定理 4.2.1 ），我們知道 Hessian 矩陣在 $a$ 是對稱的。

接下來，我們考慮的函數 $f$ 都會滿足二階導數連續的假設，所以他的 Hessian 矩陣會是對稱的。

**命題 4.2.6 ：** 與定義 4.2.5 相同的假設之下，我們有

$$H_f(a) = J_f(\overrightarrow{\operatorname{grad}} f(a))^T.$$

**證明：**這是可以直接透過把 Jacobi 矩陣的定義，用在梯度向量上所得到的。 □

在好的假設之下（二階導數的連續性），當我們討論函數 $f: A \subseteq \mathbb{R}^n \to \mathbb{R}$ 的局部行為時，可以使用 Hessian 矩陣的對稱性，討論由他所定義出來的二次型 (quadratic form) 的性質。更確切來說，二次型在臨界點的性質，可以告訴我們這個臨界點會是局部最大值，局部最小值，還是個鞍點 (saddle point)。我們在後面的第 4.3.2 小節還有第 4.3.3 小節會有更多討論。

## 第三小節 高階微分

給定從一個賦範空間 $V$ 中的開集 $A$ 映射至另一個賦範空間 $W$ 的函數 $f: A \subseteq V \to W$，在存在性的假設之下，我們在定義 4.1.1 定義了他在點 $a \in A$ 的微分，還有在定義 4.1.3 中定義了他的微分映射 $Df$。如果我們把微分映射 $Df: A \to \mathcal{L}_c(V, W)$ 微分，則我們可以定義他高階微分。

從定義 4.1.1 ，我們知道 $Df$ 的微分應該取值在 $\mathcal{L}_c(V, \mathcal{L}_c(V, W))$ 中，而這個空間也可以看成由 $V \times V$ 到 $W$ 的連續雙線性映射所構成的空間 $\mathcal{L}_c^2(V \times V, W)$。我們可以由下列映射看出這個對應：

$$\begin{array}{ccc} \mathcal{L}_c(V, \mathcal{L}_c(V, W)) & \to & \mathcal{L}_c^2(V \times V, W) \\ \Phi & \mapsto & \left\{ \begin{array}{ccc} V \times V & \to & W \\ (x, y) & \mapsto & \Phi(x)(y) \end{array} \right. \end{array}.$$

相似地，$p \geqslant 1$ 階導數會取值在空間 $\mathcal{L}_c^p(V^p, W)$ 中，而他可以看作是連續 $p$ 線性映射所構成的空間。

**定義 4.2.7 ：** 我們可以透過遞迴方式來定義 $f$ 的高階微分。

- 對於 $p \geqslant 1$，如果他的 $p$ 階微分 $D^p f : A \to \mathcal{L}_c^p(V^p, W)$ 定義良好，且當 $h_{p+1} \to 0$ 時，存在 $\varphi_{p+1} \in \mathcal{L}_c^{p+1}(V^{p+1}, W)$，使得我們有

$$D^p f(a + h_{p+1})(h_1, \ldots, h_p) = D^p f(a)(h_1, \ldots, h_p) + \varphi_{p+1}(h_1, \ldots, h_p, h_{p+1}) + o(\|h_{p+1}\|_V)$$

  其中漸進行為對於 $V^p$ 中有界的 $(h_1, \ldots, h_p)$ 來說是均勻的，則我們說 $f$ 在 $a \in A$ 可以被微分 $p+1$ 次。如果這樣的映射 $\varphi_{p+1}$ 存在，則他會是唯一的，稱作 $f$ 在 $a$ 的 $(p+1)$ 階微分，記作 $D^{p+1} f(a)$。

- 對於 $p \geqslant 1$，如果 $D^p f$ 在 $A$ 上定義良好且連續，則我們說 $f$ 是 $\mathcal{C}^p$ 類的。

**註解 4.2.8 ：** 如果我們取 $V = \mathbb{R}^n$，$W = \mathbb{R}$，以及 $A \subseteq V$ 為開子集合。考慮 $\mathcal{C}^1$ 函數 $f : A \to W$，並假設他的二階導數存在。固定 $a \in A$ 並取 $\varepsilon > 0$ 使得 $B(a, \varepsilon) \subseteq A$。那麼，對於 $h_2 \in B(0, \varepsilon)$，我們有

$$\begin{aligned}
Df(a + h_2)(h_1) &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a + h_2)\, \mathrm{d}x_i(h_1) \\
&= \sum_{i=1}^n \Big[ \frac{\partial f}{\partial x_i}(a) + \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j \partial x_i}(a) h_{2,j} + o(\|h_2\|_V) \Big] \mathrm{d}x_i(h_1) \\
&= Df(a)(h_1) + \sum_{i=1}^n \sum_{j=1}^n h_{2,j} \frac{\partial^2 f}{\partial x_j \partial x_i}(a) h_{1,i} + o(\|h_2\|_V)\, O(\|h_1\|_V).
\end{aligned}$$

這蘊含 $D^2 f(a)$ 是由 Hessian $H_f(a)$ 給出的（連續）雙線性泛函，寫做

$$(h_1, h_2) \mapsto \sum_{i=1}^n \sum_{j=1}^n h_{2,j} \frac{\partial^2 f}{\partial x_j \partial x_i}(a) h_{1,i} = h_2^T H_f(a) h_1.$$

介於高階微分與高階導數之間，我們也有類似的關係式，但這裡我們不多做探討，因為這不是這門課的目的。

在下面的小節，我們會考慮相同的架構，也就是 $V = \mathbb{R}^n$ 以及 $W = \mathbb{R}$，並考慮定義在開子集合 $A \subseteq V$ 上函數 $f : A \to W$ 的泰勒展開式。在這個情況中，我們只需要高階微分 $D^p f$ 取值在 $(\underbrace{h, \ldots, h}_{p\text{ 個}})$ 的情況。

## 第三節 實函數的局部性質

在這個章節中，我們會討論實函數的局部行為。

## 第一小節 Taylor 展開式

令 $p \geqslant 1$ 為整數。我們回顧對於 $\mathcal{C}^p$ 類的函數 $f : I \to \mathbb{R}$，其中 $I \subseteq \mathbb{R}$ 是個開區間，我們會有下面的 Taylor 展開式。令 $x \in I$ 以及 $h \in \mathbb{R}$ 使得 $x + h \in I$。

$$\begin{aligned}
&\text{Taylor–Lagrange} && f(x+h) = f(x) + \sum_{m=1}^{p-1} f^{(m)}(x)\frac{h^m}{m!} + f^{(p)}(c)\frac{h^p}{p!} \text{ 對於某個 } c \in (x, x+h).\\
&\text{Taylor integral} && f(x+h) = f(x) + \sum_{m=1}^{p-1} f^{(m)}(x)\frac{h^m}{m!} + h^p \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} f^{(p)}(x+th)\,\mathrm{d}t.\\
&\text{Taylor–Young} && f(x+h) = f(x) + \sum_{m=1}^{p} f^{(m)}(x)\frac{h^m}{m!} + o(|h|^p) \text{ 當 } h \to 0.
\end{aligned}$$

接下來，我們會把這些展開式推廣到定義在高維度歐氏空間 $\mathbb{R}^n$ 上的實函數。

令 $A$ 為 $\mathbb{R}^n$ 中的開子集合，$f : A \to \mathbb{R}$ 是個 $\mathcal{C}^p$ 類函數，其中 $p \geqslant 1$，還有 $a \in A$。在定義 4.1.1 中，我們已經定義過 $f$ 在 $a$ 的微分 $\mathrm{d}f_a$，而且在命題 4.1.17 中，我們也給出了微分與方向微分 $f'_u(a)$ 之間的關係式。此外，從定理 4.1.21 我們得知，這也能使用 $f$ 在 $a$ 的偏微分來改寫。接下來，我們會定義 $f$ 的高階方向微分。

我們會看到，高維度的泰勒展開式與他們一維的版本其實相差不大，因為當我們把函數 $f$ 限制在線段 $[x, x+h]$ 上面時，我們會得到一個在一維子空間上的函數。

**定義 4.3.1 ：** 對於 $1 \leqslant m \leqslant p$，我們把 $f$ 在 $a$ 沿著方向 $u \in \mathbb{R}^n$ 的 $m$ 階微分定義做

$$\begin{aligned}
f_u^{(m)}(a) &= \sum_{i_m=1}^{n} \cdots \sum_{i_1=1}^{n} \frac{\partial^m f}{\partial x_{i_m} \ldots \partial x_{i_1}}(a) u_{i_1} \ldots u_{i_m}, && (4.19)\\
&= \sum_{j_1+\cdots+j_n=m} \frac{m!}{j_1! \ldots j_n!} \frac{\partial^m f}{\partial x_1^{j_1} \ldots \partial x_n^{j_n}}(a) u_1^{j_1} \ldots u_n^{j_n}, && (4.20)
\end{aligned}$$

其中的等式是可以由定理 4.2.1 直接得到的。

**定理 4.3.2 【Taylor–Lagrange 展開式】：** 令 $x \in A$ 以及 $h \in \mathbb{R}^n$ 使得 $[x, x+h] \subseteq A$。那麼存在 $t \in (0, 1)$ 使得

$$f(x+h) = f(x) + \sum_{m=1}^{p-1} \frac{f_h^{(m)}(x)}{m!} + \frac{f_h^{(p)}(x+th)}{p!}. \tag{4.21}$$

**證明：** 由於 $A$ 是開集且 $[x, x+h] \subseteq A$，存在 $\delta > 0$ 使得 $x + th \in A$ 對於所有 $t \in (-\delta, 1+\delta)$。令 $g : (-\delta, 1+\delta) \to \mathbb{R}$ 定義做

$$g(t) = f(x+th), \quad \forall t \in (-\delta, 1+\delta). \tag{4.22}$$

我們注意到，$g$ 還是個 $\mathcal{C}^p$ 的函數，因為他是這種函數的合成。我們也會有 $f(x+h) - f(x) = g(1) - g(0)$。我們可以把一維的 Taylor 展開式用在 $g$ 上，也就是說

$$g(1) - g(0) = \sum_{m=1}^{p-1} \frac{g^{(m)}(0)}{m!} + \frac{g^{(p)}(t)}{p!} \quad \text{對於某個 } t \in (0,1).$$

透過鏈鎖律（命題 4.1.9），我們可以把 $g$ 的微分寫下來。對於 $t \in (-\delta, 1+\delta)$，我們有

$$g'(t) = \mathrm{d}f_{x+th}(h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x+th)h_i = f'_h(x+th),$$
$$g''(t) = \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(x+th)h_i h_j = f_h^{(2)}(x+th).$$

再使用遞迴，我們不難得到

$$g^{(m)}(t) = f_h^{(m)}(x+th), \quad \text{以及} \quad g^{(m)}(0) = f_h^{(m)}(x), \quad \forall m \geqslant 1.$$

這就是我們所想要證明的。 □

透過考慮函數 $g$，並使用與式 (4.22) 中相同的技巧，搭配上其他的一維泰勒展開式，我們可以輕易得到下列其他展開式。

**定理 4.3.3** 【Taylor 展開式與積分餘項】：令 $x \in A$ 以及 $h \in \mathbb{R}^n$ 使得 $[x, x+h] \subseteq A$。那麼我們有

$$f(x+h) = f(x) + \sum_{m=1}^{p-1} \frac{f_h^{(m)}(x)}{m!} + \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} f_h^{(p)}(x+th)\,\mathrm{d}t. \tag{4.23}$$

**證明：** 見習題 4.19。 □

**定理 4.3.4** 【Taylor-Young 展開式】：令 $x \in A$。那麼，當 $h \to 0$ 時，我們有

$$f(x+h) = f(x) + \sum_{m=1}^{p} \frac{f_h^{(m)}(x)}{m!} + o(|h|^p). \tag{4.24}$$

**證明：**見習題 4.19 。 □

## 第二小節 二次型

**定義 4.3.5 ：** 給定對稱矩陣 $A \in \mathcal{M}_n(\mathbb{R})$，我們可以定義 $\mathbb{R}^n$ 上的二次型 (quadratic form)：

$$q_A(x) = q_A(x_1, \ldots, x_n) = \sum_{1 \leqslant i,j \leqslant n} a_{ij} x_i x_j = x^T A x, \quad \forall x \in \mathbb{R}^n, \tag{4.25}$$

在上式中，在 $\mathbb{R}^n$ 中的向量可以被視為是個行向量。

**定義 4.3.6 ：** 給定二次型 $Q$，如果

- 對於所有 $x \in \mathbb{R}^n$，我們有 $Q(x) \geqslant 0$，則我們說他是正的；
- 對於所有 $x \in \mathbb{R}^n \backslash \{0\}$，我們有 $Q(x) > 0$，則我們說他是正定的；
- 對於所有 $x \in \mathbb{R}^n$，我們有 $Q(x) \leqslant 0$，則我們說他是負的；
- 對於所有 $x \in \mathbb{R}^n \backslash \{0\}$，我們有 $Q(x) < 0$，則我們說他是負定的。

**註解 4.3.7 ：** 在二階偏微分都在 $a \in A$ 連續的假設之下，我們可以把 $f_u^{(2)}(a)$ 改寫做：

$$f_u^{(2)}(a) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} u_i u_j = u^T H_f(a) u,$$

其中 $H_f(a)$ 是個對稱矩陣，且向量 $u$ 可以被視為行向量，且 $u^T$ 是他的轉置向量。這是在定義 4.3.5 中提到的二次型。

**註解 4.3.8 ：** 從線性代數的課程中，我們知道對稱矩陣 $A$ 是可對角化的；也就是說，我們能找到對角矩陣 $D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$ 滿足 $\lambda_1 \geqslant \ldots \geqslant \lambda_n$，以及正交矩陣 $P$（也就是說滿足 $PP^T = P^T P = I_n$）使得 $A = P^T D P$。這代表著，根據適當由 $P$ 給出的基底變換之後，二次型會是對角的。更確切來說，令 $v = Pu$，則我們有

$$u^T A u = (Pu)^T D (Pu) = v^T D v,$$

這代表著

$$q_A(u) = q_D(v) = \sum_{i=1}^n \lambda_i |v_i|^2.$$

因此，我們可以總結，如果 $\lambda_n > 0$，那麼二次型是正定的；如果 $\lambda_1 < 0$，那麼二次型是負定的。

## 第三小節 局部極值

接下來，令 $A$ 為 $\mathbb{R}^n$ 的子集合，且 $f: A \to \mathbb{R}$ 為函數。我們想要討論 $f$ 的局部極值。我們會使用在第 4.3.1 小節中所得到的 Taylor 展開式。

**定義 4.3.9 ：** 如果 $f$ 在內點 $a \in \mathring{A}$ 是可微的，且 $\mathrm{d}f_a = 0$，則我們稱 $a$ 為 $f$ 的臨界點。

**命題 4.3.10 ：** 假設 $f$ 在內點 $a \in \mathring{A}$ 是個局部極值，且 $f$ 在 $a$ 可微，那麼 $a$ 是個 $f$ 的臨界點。

**證明：** 不失一般性，我們可以假設 $f$ 在 $a$ 是個局部最大值。令 $h \in \mathbb{R}^n$，我們想要證明 $\mathrm{d}f_a(h) = 0$。由於 $a \in \mathring{A}$，那麼會存在 $\eta > 0$ 使得 $[a - \eta h, a + \eta h] \subseteq A$。我們定義映射 $\varphi: [-\eta, \eta] \to \mathbb{R}, t \mapsto f(a + th)$，他在 $t = 0$ 有局部最大值。由於 $f$ 在 $a$ 可微，我們知道 $\varphi$ 在 $0$ 可微，且我們有 $\varphi'(0) = \mathrm{d}f_a(h)$。此外，我們還有

$$\varphi'(0) = \lim_{\substack{t \to 0 \\ t > 0}} \frac{\varphi(t) - \varphi(0)}{t} \leqslant 0, \quad \text{以及} \quad \varphi'(0) = \lim_{\substack{t \to 0 \\ t < 0}} \frac{\varphi(t) - \varphi(0)}{t} \geqslant 0,$$

這會讓我們得到 $\varphi'(0) = 0$。 □

**註解 4.3.11 ：** 命題 4.3.10 告訴我們，如果要找函數 $f: A \to \mathbb{R}$ 的局部極值，我們需要去考慮下面這些點：

(i) $a \in \mathring{A}$ 是 $f$ 的臨界點；

(ii) $a \in \mathring{A}$ 使得 $f$ 在 $a$ 不可微；

(iii) $a \in A \backslash \mathring{A}$。

**定理 4.3.12 ：** 假設 $f$ 是 $\mathcal{C}^2$ 類的，且存在 $a \in A$ 滿足 $\mathrm{d}f_a = 0$。Taylor-Young 展開式（式 (4.24)）告訴我們

$$f(a + h) = f(a) + \frac{1}{2} Q(h) + o(\|h\|^2), \quad \text{當 } h \to 0.$$

(1) 如果 $f$ 在 $a$ 是個局部最小值（局部最大值），那麼 $Q$ 是個正二次型（負二次型）。

(2) 如果 $Q$ 是正定二次型（負定二次型），那麼 $f$ 在 $a$ 是個局部最小值（局部最大值）。

**範例 4.3.13 ：** 在定理 4.3.12 (2) 中，如果二次型只是正的，並不足以得到局部最小值。我們可以考慮函數 $f: \mathbb{R} \to \mathbb{R}, x \mapsto x^3$ 在 $a = 0$ 的情況，二次型 $Q = 0$ 但 $f$ 並沒有局部極值。

**證明：**

(1) 假設 $f$ 在 $a$ 是個局部最小值。令 $h \in \mathbb{R}^n$ 以及 $t \in \mathbb{R}$。當 $t$ 夠靠近 0 時，我們有

$$f(a+th) = f(a) + \frac{1}{2}Q(th) + o(\|th\|^2) \geqslant f(a).$$

這讓我們得到

$$0 \leqslant Q(th) + o(\|th\|^2) = t^2(Q(h) + o(1)),$$

也就是說，當我們取 $t \to 0$ 時，可以推得 $Q(h) \geqslant 0$。

(2) 假設 $Q$ 是正定二次型，那麼對於 $h \in \mathbb{R}^n$ 還有 $h \neq 0$，我們會有 $Q(h) > 0$。由於 $\mathbb{R}^n$ 的單位球殼 $S(0,1)$ 是緊緻的，我們能推得 $m = \inf_{h \in S(0,1)} Q(h) > 0$。因此，當 $h \to 0$ 時，我們有

$$f(a+h) - f(h) = \frac{1}{2}[Q(h) + o(\|h\|^2)] = \frac{\|h\|^2}{2}\Big[Q\Big(\frac{h}{\|h\|}\Big) + o(1)\Big] \geqslant \frac{\|h\|^2}{2}(m + o(1)).$$

當 $h$ 夠靠近 0 時，我們有 $m + o(1) \geqslant 0$，因此可以推得 $f(a+h) \geqslant f(a)$。 □

**範例 4.3.14 ：** 這裡我們取 $n = 2$ 的情況當例子。$\mathbb{R}^2$ 上的二次型可以被下面這個對稱矩陣所表示：

$$A = \begin{pmatrix} r & s \\ s & t \end{pmatrix} \in \mathcal{M}_2(\mathbb{R}).$$

根據註解 4.3.8 ，我們知道 $A = P^T D P$，其中 $P$ 是個正交矩陣，且 $D = \mathrm{Diag}(\lambda_1, \lambda_2)$ 是個對角矩陣，滿足 $\lambda_1 \geqslant \lambda_2$。這讓我們能得到下面與特徵值 $\lambda_1 \geqslant \lambda_2$ 相關的關係式：

$$\begin{cases} \lambda_1 + \lambda_2 = \mathrm{tr}(D) = \mathrm{tr}(A) = r + t, \\ \lambda_1 \lambda_2 = \det(D) = \det(A) = rt - s^2. \end{cases}$$

因此，我們可以討論下面情況：

(i) 當 $rt - s^2 > 0$ 還有 $r + t > 0$ 時，由 $A$ 定義出來的二次型是正定的。

(ii) 當 $rt - s^2 > 0$ 還有 $r + t < 0$ 時，由 $A$ 定義出來的二次型是負定的。

當我們把這個結果應用在 $\mathcal{C}^2$ 類的函數 $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ 上時，考慮 $f$ 的臨界點 $a \in \mathring{A}$。我們記

$$r = \frac{\partial^2 f}{\partial x^2}(a), \quad s = \frac{\partial^2 f}{\partial x \partial y}(a), \quad t = \frac{\partial^2 f}{\partial y^2}(a).$$

那麼，從上面的討論，我們得知：

(i) 當 $rt - s^2 > 0$ 還有 $r + t > 0$ 時，$f$ 在 $a$ 點有局部最小值。

(ii) 當 $rt - s^2 > 0$ 還有 $r + t < 0$ 時，$f$ 在 $a$ 點有局部最大值。

(iii) 當 $rt - s^2 < 0$ 時，$f$ 在 $a$ 點沒有局部極值，且我們稱他為鞍點 (saddle point)。

(iv) 當 $rt - s^2 = 0$ 時，我們無法總結。

# 第四節 隱函數定理

## 第一小節 反函數定理

對於 $\mathcal{C}^1$ 類函數 $f : \mathbb{R} \to \mathbb{R}$，我們知道如果對於所有 $x \in \mathbb{R}$，我們有 $f'(x) \neq 0$，那麼 $f$ 是個雙射函數，且他的反函數 $f^{-1}$ 也是個 $\mathcal{C}^1$ 類函數，且滿足 $(f^{-1})'[f(x)] = [f'(x)]^{-1}$ 對於所有 $x \in \mathbb{R}$。

令 $V$ 與 $W$ 為兩個 Banach 空間，且 $A \subseteq V$ 是個 $V$ 的開子集合。

**定理 4.4.1**【局部反函數定理】：令 $f : A \to W$ 為 $\mathcal{C}^1$ 類函數。假設存在 $a \in A$ 使得 $(\mathrm{d}f_a)^{-1}$ 存在且 $\mathrm{d}f_a$ 和 $(\mathrm{d}f_a)^{-1}$ 皆連續（我們說 $\mathrm{d}f_a$ 是個雙連續同構變換）。那麼會存在包含 $a$ 的開集合 $X$ 以及包含 $f(a)$ 的開集合 $Y$ 使得

(i) 函數 $f_{|X}$ 是個介於 $X$ 與 $Y$ 之間的雙射函數；

(ii) 反函數 $g := (f_{|X})^{-1} : Y \to X$ 是連續的；

(iii) $g$ 是 $\mathcal{C}^1$ 類的且對於所有 $x \in X$，我們有 $\mathrm{d}g_{f(x)} = (\mathrm{d}f_x)^{-1}$。

在這個情況中，我們也說 $f_{|X} : X \to Y$ 是個 $X$ 與 $Y$ 之間的 $\mathcal{C}^1$ 微分同胚變換，或是 $f : A \to W$ 在 $a$ 附近是個局部 $\mathcal{C}^1$ 微分同胚變換。

**註解 4.4.2**：

(1) 這被稱做局部反函數定理，因為他只描述了在 $a \in X$ 附近以及 $f(a) \in Y$ 附近的局部行為。稍後在系理 4.4.5 中，我們會看到怎麼把他升級為全域反函數定理。

(2) 如果我們考慮 $n \geqslant 1$ 以及 $V = W = \mathbb{R}^n$，由於 $\mathcal{L}(V, W) = \mathcal{L}_c(V, W)$，局部反函數定理在 $a \in A \subseteq V$ 所要求的條件會簡化為 $\mathrm{d}f_a$ 可逆，也就是說 $\det J_f(a) \neq 0$。

**範例 4.4.3**：

(1) 我們考慮 $\mathbb{R}$ 上的 $\mathcal{C}^1$ 類函數 $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$。對於 $a \in \mathbb{R}\backslash\{0\}$，他的微分寫做 $f'(a) = 2a \neq 0$，根據局部反函數定理，當我們把 $f$ 限制在一個包含 $a$ 的開集 $X$ 上時，他的反函數是定義良好的。事實上，當 $a > 0$，我們可以取 $X = Y = (0, \infty)$，並定義

$g(y) = \sqrt{y}$ 對於 $y \in Y$；當 $a < 0$ 時，我們可以取 $X = (-\infty, 0), Y = (0, \infty)$，並定義 $g(y) = -\sqrt{y}$ 對於 $y \in Y$。

(2) 如果我們定義極座標與卡式座標之間的轉換：

$$\begin{array}{rccc} \varphi: & (0,\infty)\times\mathbb{R} \subseteq \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (r,t) & \mapsto & (r\cos t, r\sin t), \end{array}$$

那麼他在 $(r,t) \in (0,\infty)\times\mathbb{R}$ 的微分寫做：

$$\mathrm{d}\varphi_{r,t}(r',t') = (r'\cos t - t'r\sin t, r'\sin t + t'r\cos t) = \begin{pmatrix} \cos t & -r\sin t \\ \sin t & r\cos t \end{pmatrix}\begin{pmatrix} r' \\ t' \end{pmatrix}.$$

這給我們

$$\det J_\varphi(r,t) = r \neq 0, \quad \text{其中 } J_\varphi(r,t) = \begin{pmatrix} \cos t & -r\sin t \\ \sin t & r\cos t \end{pmatrix}.$$

根據局部反函數定理，在每個 $(r,t) \in (0,\infty)\times\mathbb{R}$，我們可以找到包含 $(r,t)$ 的開集 $X$ 使得 $f$ 在 $X$ 上是可逆的。然而，$f$ 並沒有全域的反函數，因為他顯然不是單射的。

**證明：**不失一般性，我們可以考慮函數 $x \mapsto (\mathrm{d}f_a)^{-1}[f(a+x) - f(a)]$，並把他記作 $f$，這樣一來，我們可以假設 $a = 0$、$f(a) = 0$，且 $\mathrm{d}f_0 = \mathrm{d}f_a = \mathrm{id}_V$，所以 $V = W$。使用 $f$ 是 $\mathcal{C}^1$ 類的假設，存在 $r > 0$ 使得

$$\overline{B}(0,r) \subseteq A \quad \text{且} \quad |||\mathrm{d}f_x - \mathrm{d}f_0||| = |||\mathrm{d}f_x - \mathrm{id}_V||| \leqslant \tfrac{1}{2}, \quad \forall x \in B(0,r).$$

那麼，對於 $x \in B(0,r)$，我們有 $\mathrm{d}f_x = \mathrm{id}_V - u$，其中 $u = \mathrm{id}_V - \mathrm{d}f_x$ 滿足 $|||u||| \leqslant \frac{1}{2}$，再根據命題 3.2.20 ，我們會有

$$\begin{aligned} (\mathrm{d}f_x)^{-1} &= \mathrm{id}_V + \sum_{n\geqslant 1} u^n, \\ \left|\left|\left|(\mathrm{d}f_x)^{-1}\right|\right|\right| &\leqslant \sum_{n\geqslant 0} |||u|||^n \leqslant 2. \end{aligned} \tag{4.26}$$

(i) 首先，讓我們來證明 $f$ 有局部的反函數。更確切來說，我們想要證明對於每個 $y \in B(0,\frac{r}{2})$，存在唯一的 $x \in B(0,r)$ 滿足 $f(x) = y$。我們會使用固定點定理（定理 3.2.7 ）來構造這樣的函數。

令 $y \in B(0,\frac{r}{2})$ 並考慮函數

$$\begin{array}{rccc} h: & B(0,r) & \to & V \\ & x & \mapsto & y + x - f(x). \end{array}$$

函數 $h$ 是 $\mathcal{C}^1$ 類的，且對於每個 $x \in B(0,r)$，我們有 $|||\mathrm{d}h_x||| = |||\mathrm{id}_V - \mathrm{d}f_x||| \leqslant \frac{1}{2}$。因此，

根據均值不等式（定理 4.1.13），我們有

$$\forall x, x' \in \overline{B}(0,r), \quad \|h(x) - h(x')\| \leqslant \frac{1}{2}\|x - x'\|. \tag{4.27}$$

因此，對於 $x \in \overline{B}(0,r)$，我們有

$$\|h(x)\| \leqslant \|y\| + \|x - f(x)\| = \|y\| + \|h(x) - h(0)\| \leqslant \|y\| + \frac{1}{2}\|x\| < r.$$

這代表著，$h$ 是個從 $\overline{B}(0,r)$ 到 $B(0,r) \subseteq \overline{B}(0,r)$ 的收縮函數，因此我們可以使用固定點定理（定理 3.2.7），進而得到存在唯一的 $x \in \overline{B}(0,r)$ 使得 $h(x) = x$。但由於 $h$ 取值在 $B(0,r)$ 中，我們知道這個固定點 $x$ 會在 $B(0,r)$ 中，且我們有 $f(x) = y$。

最後，我們令 $Y = B(0, \frac{r}{2})$ 以及 $X = f^{-1}(Y) \cap B(0,r)$ 來總結。根據 $f$ 的連續性以及 $f(0) = 0$，開集 $X$ 也包含 $0$。再根據我們上面所證明的，限制函數 $f_{|X} : X \to Y$ 是個雙射函數。

(ii) 令 $g : Y \to X$ 為 $f_{|X}$ 的反函數，也就是 $g = (f_{|X})^{-1}$。考慮函數 $h : X \to V, x \mapsto x - f(x)$，所以對於所有 $x \in X$，我們有 $x = f(x) + h(x)$。那麼，對於 $x, x' \in B(0,r)$，我們有

$$\begin{aligned} & \|x - x'\| \leqslant \|h(x) - h(x')\| + \|f(x) - f(x')\| \leqslant \frac{1}{2}\|x - x'\| + \|f(x) - f(x')\| \\ \Leftrightarrow \quad & \|x - x'\| \leqslant 2\|f(x) - f(x')\|. \end{aligned}$$

因此，對於 $y, y' \in Y$，我們有

$$\|g(y) - g(y')\| \leqslant 2\|f(g(y)) - f(g(y'))\| = 2\|y - y'\|. \tag{4.28}$$

這告訴我們 $g$ 是個 Lipschitz 函數，所以也是連續的。

(iii) 令 $x \in X$ 以及 $y = f(x) \in Y$。讓我們先來檢查 $\mathrm{d}g_y = (\mathrm{d}f_x)^{-1}$。令 $w \in W$ 使得 $y+w \in Y$，令 $v = g(y+w) - g(y)$，這會與 $w = f(x+v) - f(x)$ 等價。根據式 (4.28)，我們有 $\|v\| \leqslant 2\|w\|$。令

$$\begin{aligned} \Delta(w) &= g(y+w) - g(y) - (\mathrm{d}f_x)^{-1}(w) \\ &= (\mathrm{d}f_x)^{-1} \circ \mathrm{d}f_x(v) - (\mathrm{d}f_x)^{-1}[f(x+v) - f(x)] \\ &= -(\mathrm{d}f_x)^{-1}[f(x+v) - f(x) - \mathrm{d}f_x(v)]. \end{aligned}$$

從式 (4.26)，我們推得

$$\|\Delta(w)\| \leqslant 2\|f(x+v) - f(x) - \mathrm{d}f_x(v)\| = 2\|v\|\,\varepsilon(v),$$

其中函數 $\varepsilon$ 滿足 $\lim_{v\to 0}\varepsilon(v) = 0$。令 $\widetilde{\varepsilon}(w) = \varepsilon\big(g(y+w) - g(y)\big)$。由於 $g$ 是連續的，我們也會有 $\lim_{w\to 0}\widetilde{\varepsilon}(w) = 0$。因此

$$\frac{\|\Delta(w)\|}{\|w\|} \leqslant \frac{2\|v\|}{\|w\|}\widetilde{\varepsilon}(w) \xrightarrow[w\to 0]{} 0.$$

這代表著 $g$ 在 $y$ 可微，且微分寫做 $\mathrm{d}g_y = (\mathrm{d}f_x)^{-1}$。

最後我們總結。由於 $u \mapsto u^{-1}$ 在可逆自同態中是個連續映射（範例 4.1.6 以及命題 4.1.7），且 $g$ 是連續的，我們推得 $y \mapsto \mathrm{d}g_y = (\mathrm{d}f_{g(y)})^{-1}$ 也是連續的，也就是說 $g$ 是個 $\mathcal{C}^1$ 類的函數。

□

**系理 4.4.4 ：** 令 $f : A \to W$ 為 $\mathcal{C}^1$ 函數。假設對於所有 $x \in A$，$\mathrm{d}f_x$ 是可逆且雙連續的。那麼，$f$ 是個開函數，也就是說，對於任意開子集合 $X \subseteq A$，像 $f(X)$ 在 $W$ 中是個開集。

**證明：**我們只需要證明 $X = A$ 的情況。對於每個 $a \in A$，局部反函數定理（定理 4.4.1）給我們包含 $a$ 的開子集合 $X_a$ 以及包含 $f(a)$ 的開子集合 $Y_a$ 使得 $f_{|X_a}$ 是個介於 $X_a$ 與 $Y_a$ 之間的雙射函數，也就是說 $f(X_a) = Y_a$。因此，

$$f(A) = f\Big(\bigcup_{a\in A} X_a\Big) = \bigcup_{a\in A} f(X_a) = \bigcup_{a\in A} Y_a,$$

所以也還是個 $W$ 中的開子集合。 □

**系理 4.4.5 【全域反函數定理】：** 令 $f : A \to W$ 為 $\mathcal{C}^1$ 類的單射函數。那麼，下列性質是等價的。

(a) 對於所有 $a \in A$，微分 $\mathrm{d}f_a$ 是可逆且雙連續的。

(b) $B = f(A)$ 在 $W$ 中是個開集，且 $f^{-1} : B \to A$ 是個 $\mathcal{C}^1$ 類函數。

如果上面其中一個性質成立，我們說 $f : A \to B$ 是個介於 $A$ 與 $B$ 之間的 $\mathcal{C}^1$ 微分同胚變換。

**證明：**

- (a) ⇒ (b). 從系理 4.4.4 ，我們知道 $B = f(A)$ 是個開集。由於 $f$ 是單射的，我們推得介於開集 $A$ 與開集 $B$ 之間的函數 $f$ 是雙射的。接著，我們需要檢查 $f^{-1}$ 是 $\mathcal{C}^1$ 類的。令 $x \in A$ 以及 $y = f(x) \in B$。局部反函數定理（定理 4.4.1）讓我們能找到包含 $x$ 的開集 $A_x$ 以及包含 $f(x)$ 的開集 $B_x$ 使得 $f_{|A_x} : A_x \to B_x$ 是雙射的，且 $(f_{|A_x})^{-1}$ 是 $\mathcal{C}^1$ 類函數。由於 $(f^{-1})_{|B} = (f_{|A})^{-1}$、$(f^{-1})_{|B_x} = (f_{|A_x})^{-1}$，且 $\mathcal{C}^1$ 類是個局部性質，我們知道 $f^{-1}$ 在 $f(x)$ 附近也是 $\mathcal{C}^1$ 類的。這對於所有 $x \in A$ 皆成立，所以 $f^{-1}$ 在 $B$ 上是 $\mathcal{C}^1$ 類的。
- (b) ⇒ (a). 記 $g = f^{-1}$。由於 $f$ 和 $g$ 都是 $\mathcal{C}^1$ 類的，透過等式 $g \circ f = \mathrm{Id}_A$ 以及鏈鎖律，我們得知對於 $x \in A$，我們有 $\mathrm{d}g_{f(x)} \circ \mathrm{d}f_x = \mathrm{Id}_V$。相似地，$f \circ g = \mathrm{Id}_B$ 還告訴我們對於所有 $x \in A$，我們有 $\mathrm{d}f_x \circ \mathrm{d}g_{f(x)} = \mathrm{Id}_W$。因此，對於所有 $x \in A$，微分 $\mathrm{d}f_x$ 是可逆的，且他的

反函數寫做 $\mathrm{d}g_{f(x)}$，而且這是個連續函數。 □

**註解 4.4.6 ：** 我們可以做個與註解 4.4.2 (2) 中類似的觀察。如果我們考慮 $n \geqslant 1$ 以及歐氏空間 $V = W = \mathbb{R}^n$，由於 $\mathcal{L}(V, W) = \mathcal{L}_c(V, W)$，我們可以把性質 (a) 改成：

(a') $\mathrm{d}f_a$ 是可逆的，或是 $\det J_f(a) \neq 0$，

並不需要要求雙連續性。

## 第二小節 微分同胚

**定義 4.4.7 ：** 令 $V, W$ 為兩個賦範向量空間，且 $A \subseteq V$ 及 $B \subseteq W$ 為開子集合。對於 $k \geqslant 1$，如果函數 $f : A \to B$ 是雙射的，且 $f$ 與 $f^{-1}$ 皆是 $\mathcal{C}^k$ 類函數，則我們說 $f$ 是個 $\mathcal{C}^k$ 類的微分同胚。

下面兩個系理給出在歐氏空間的情況中，什麼時候我們的函數會是局部微分同胚或是全域微分同胚。他們的證明是建立在局部反函數定理還有全域反函數定理。我們注意到，如果我們加上雙連續性的假設，那麼我們可以把這些證明推廣到一般的 Banach 空間中。由於在一般的 Banach 空間或是賦範向量空間中，在 $k \geqslant 2$ 的情況，我們沒有對 $\mathcal{C}^k$ 函數多做討論（見第 4.2.3 小節），因此我們把下面的敘述侷限在歐氏空間中，因為在第 4.2.1 小節中，我們有討論過這種空間中的規律性。

**系理 4.4.8 ：** 令 $A \subseteq \mathbb{R}^n$ 為開子集合，以及 $f : A \to \mathbb{R}^n$ 為 $\mathcal{C}^k$ 類函數。假設存在 $a \in A$ 使得 $\mathrm{d}f_a$ 是可逆的（或 $\det J_f(a) \neq 0$），那麼存在包含 $a$ 的開集合 $X_a$ 以及包含 $f(a)$ 的開集合 $Y_a$ 使得 $f_{|X_a}$ 是個從 $X_a$ 到 $Y_a$ 的 $\mathcal{C}^k$ 微分同胚。此外，對於所有 $x \in X_a$，我們也有 $\mathrm{d}(f_{|X_a}^{-1})_{f(x)} = (\mathrm{d}f_x)^{-1}$。

**證明：** 由於 $\mathrm{d}f_a$ 是可逆的，且我們在有限維度向量空間中，$\mathrm{d}f_a$ 自動會是雙連續的。接著，我們可以使用局部反函數定理（定理 4.4.1 ）來找到 $X_a$ 以及 $Y_a$，使得 $f_{|X_a}$ 是個 $\mathcal{C}^1$ 微分同胚。再來我們只需要檢查 $g = f_{|X_a}^{-1}$ 是個 $\mathcal{C}^k$ 函數即可。

我們回顧記號 $J_f(a)$ 的意義，他代表的是 $f$ 在 $a$ 的 Jacobi 矩陣，且 $J_g(f(a))$ 是 $g$ 在 $f(a)$ 的 Jacobi 矩陣。對於所有 $x \in X_a$，由於 $(\mathrm{d}g)_{f(x)} = (\mathrm{d}f_x)^{-1}$，我們可以推得 $J_g(f(x)) = J_f(x)^{-1} = (\det J_f(x))^{-1}\widetilde{J}(x)$，其中 $\widetilde{J}(x)$ 是 $J_f(x)$ 餘子矩陣的轉置矩陣（也稱作伴隨矩陣），裡面的係數是 $J_f(x)$ 中係數乘積的線性組合。因此，$g$ 的一階偏微分會是 $f$ 一階偏微分所構成的有理函數，他們是 $\mathcal{C}^{k-1}$ 類的，所以 $g$ 的一階偏微分也是 $\mathcal{C}^{k-1}$ 類的。所以我們推得 $g$ 是 $\mathcal{C}^k$ 類的。 □

**系理 4.4.9**：令 $A \subseteq \mathbb{R}^n$ 是開子集合，且 $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ 是個 $\mathcal{C}^k$ 類的單射函數，其中 $k \geqslant 1$。那麼，下列性質是等價的。

(a) 對於所有 $a \in A$，微分 $\mathrm{d}f_a$ 是可逆的。

(b) $B = f(A)$ 在 $W$ 中是開集，且 $f$ 是個從 $A$ 到 $B$ 的 $\mathcal{C}^k$ 類微分同胚。

**證明**：證明與系理 4.4.5 和系理 4.4.8 類似。 □

## 第三小節 隱函數定理

我們先給隱函數定理背後的一些動機。我們給定函數 $f : A \subseteq \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ 然後想要找出他的等高線，也就是我們固定 $c \in \mathbb{R}^n$，想要找出 $x \in \mathbb{R}^m$ 和 $y \in \mathbb{R}^n$ 使得 $f(x, y) = c$。隱函數定理給出局部的充分條件，使得 $y$ 可以寫成 $x$ 的函數 $\varphi$，也就是說 $f(x, \varphi(x)) = c$。換句話說，在 $x$ 的附近，$f(x, y) = c$ 的解，可以用圖來表示。更一般來說，我們可以讓 $c$ 變成變數，我們會得到的是取決於 $x$ 和 $c$ 的函數 $\varphi$。下面定理給出更確切的敘述。

令

$$\begin{array}{rccc} f = (f_1, \ldots, f_n) : & A \subseteq \mathbb{R}^m \times \mathbb{R}^n & \to & \mathbb{R}^n \\ & (x, y) = (x_1, \ldots, x_m; y_1, \ldots, y_n) & \mapsto & f(x, y). \end{array} \tag{4.29}$$

我們可以對於變數 $x = (x_1, \ldots, x_m)$ 以及 $y = (y_1, \ldots, y_n)$，以下列方式定義在 $(a, b) \in A$ 的部份 Jacobi 矩陣還有他們的行列式（稱作部份 Jacobi 行列式）：

$$J_{f,x}(a, b) = \left[\frac{\partial f_i}{\partial x_j}(a, b)\right]_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant m}} \qquad \text{以及} \qquad J_{f,y}(a, b) = \left[\frac{\partial f_i}{\partial y_j}(a, b)\right]_{1 \leqslant i, j \leqslant n}.$$

**定理 4.4.10**【隱函數定理】：令 $m, n \geqslant 1$ 為整數且 $A \subseteq \mathbb{R}^m \times \mathbb{R}^n$ 為開子集合。如同在式 (4.29) 中，令 $k \geqslant 1$，且我們給定 $\mathcal{C}^k$ 類函數 $f$。固定 $(a, b) \in A$。如果部份 Jacobi 行列式 $\det J_{f,y}(a, b)$ 是非零的，那麼存在

- 包含 $a$ 的開子集合 $X$，包含 $f(a, b)$ 的開子集合 $W$，以及包含 $(a, b)$ 的開子集合 $Z$；
- 一個 $\mathcal{C}^k$ 類函數 $\varphi : X \times W \to \mathbb{R}^n$

使得對於所有 $x \in X$ 還有 $w \in W$，$y = \varphi(x, w)$ 會是 $f(x, y) = w$ 在 $(x, y) \in Z$ 條件之下唯一的解。所以，對於所有 $x \in X$ 還有 $w \in W$，我們有 $f(x, \varphi(x, w)) = w$。

此外，對於 $(a, c) \in X \times W$，我們記 $b = \varphi(a, c)$，這樣我們也會得到下列部份 Jacobi 矩陣之間的關係式：

$$J_{\varphi,x}(a, c) = -\left[J_{f,y}(a, b)\right]^{-1} J_{f,x}(a, b) \qquad \text{以及} \qquad J_{\varphi,y}(a, c) = \left[J_{f,y}(a, b)\right]^{-1}.$$

**註解 4.4.11 ：** 在前面小節中，我們給的（局部和全域）反函數定理是在一般 Banach 空間中的版本，當中我們要求微分要是可逆且雙連續的。我們也注意到，在註解 4.4.2 與註解 4.4.6 中，當我們取歐氏空間（或是有限維度賦範向量空間）時，雙連續性質自動會成立，因此不用檢查。這裡，為了簡化敘述和證明，我們給的隱函數定理是在歐氏空間中的，但要知道的是，當我們在處理一般的 Banach 空間時，唯一需要增加的條件是雙連續性的假設。

**證明：** 令 $F=(F_1,\dots,F_m;F_{m+1},\dots,F_{m+n})$ 為定義在 $A$ 上的函數，且取值在 $\mathbb{R}^{m+n}$ 中，他的分量定義如下：對 $1\leqslant i\leqslant m$，定義 $F_i(x,y)=x_i$；對 $1\leqslant i\leqslant n$，定義 $F_{m+i}=f_i(x,y)$。這樣一來，$F$ 的 Jacobi 矩陣是個分塊矩陣，寫做

$$\left(\begin{array}{c|c} I_m & \mathbf{0} \\ \hline \star & \left(\dfrac{\partial f_i}{\partial y_j}(a,b)\right)_{1\leqslant i,j\leqslant n} \end{array}\right).$$

他的行列式與部份 Jacobi 行列式 $\det J_{f,y}(a,b)$ 相同，而根據假設，這是非零的。因此，從系理 4.4.8 我們得知存在包含 $(a,b)$ 的開集 $Z$ 與包含 $F(a,b)=(a,f(a,b))$ 的開集 $Y$ 使得 $F_{|Z}$ 是個從 $Z$ 到 $Y$ 的 $\mathcal{C}^k$ 微分同胚。我們可以把 $Y$ 限制成 $X\times W\subseteq Y$，其中 $X$ 是個包含 $a$ 的開集，且 $W$ 是個包含 $f(a,b)$ 的開集。這樣一來，我們可以把 $F^{-1}:X\times W\subseteq Y\to Z$ 寫成 $F^{-1}(x,w)=(x,\varphi(x,w))$，其中 $\varphi$ 是個 $\mathcal{C}^k$ 類函數。因此，我們推得對於任何 $(x,w)\in X\times W$，會存在唯一的 $y$ 使得 $(x,y)\in Z$ 且 $f(x,y)=w$；此外，我們也有 $y=\varphi(x,z)$。

如果要得到部份 Jacobi 矩陣之間的關係，我們只需要使用命題 4.1.26 中，合成函數與 Jacobi 矩陣乘積的關係即可。 □

在上述定理中，我們可以把 $w$ 取為常數，這對應到的是下面的引理。

**系理 4.4.12 ：** 在與定理 4.4.10 相同的假設之下，我們能找到

- 包含 $a$ 的開子集合 $X$ 以及包含 $b$ 的開子集合 $Y$；
- 一個 $\mathcal{C}^k$ 類函數 $\varphi:X\to\mathbb{R}^n$

使得對於所有 $x\in X$，$y=\varphi(x)$ 會是 $f(x,y)=c$ 在條件 $y\in Y$ 之下的唯一解。這讓我們可以這樣寫：對於所有 $x\in X$，我們有 $f(x,\varphi(x))=c$。

此外，對於 $a\in X$，我們記 $b=\varphi(x)$，這樣我們會有下列部份 Jacobi 矩陣之間的關係式：

$$J_{f,x}(a,b)+J_{f,y}(a,b)J_{\varphi,x}(a)=0 \qquad \text{或} \qquad J_{\varphi,x}(a)=-\big[J_{f,y}(a,b)\big]^{-1}J_{f,x}(a,b). \tag{4.30}$$

**系理 4.4.13：** 令 $A \subseteq \mathbb{R}^2$ 為開集，$k \geqslant 1$，且 $f : \mathbb{R}^2 \to \mathbb{R}$ 是個 $\mathcal{C}^k$ 類函數。令 $(a,b) \in A$ 並假設

$$f(a,b) = 0 \quad \text{且} \quad \frac{\partial f}{\partial y}(a,b) \neq 0.$$

那麼會存在 $\alpha, \beta > 0$ 使得對於所有 $x \in (a-\alpha, a+\alpha)$，方程式 $f(x,y) = 0$ 會在 $(b-\beta, b+\beta)$ 中有唯一解 $y = \varphi(x)$。此外，函數 $\varphi$ 在 $(a-\alpha, a+\alpha)$ 上會是 $\mathcal{C}^k$ 類的，且我們有

$$\varphi'(x) = -\frac{\partial f}{\partial x}(x,\varphi(x)) \Big/ \frac{\partial f}{\partial y}(x,\varphi(x)), \quad \forall x \in (a-\alpha, a+\alpha).$$

**證明：**從系理 4.4.12 ，我們能得到 $\alpha, \beta > 0$ 的存在性以及 $\varphi$ 的規律性。要計算 $\varphi'$，我們把關係式 $f(x,\varphi(x)) = 0$ 微分，進而得到

$$\frac{\partial f}{\partial x}(x,\varphi(x)) + \varphi'(x)\frac{\partial f}{\partial y}(x,\varphi(x)) = 0.$$

這也能夠直接由式 (4.30) 推得。 □

下面這個引理也可以用類似的方式來證明。

**系理 4.4.14：** 令 $A \subseteq \mathbb{R}^3$ 為開集，$k \geqslant 1$，且 $f : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ 是個 $\mathcal{C}^k$ 類函數。令 $(a,b,c) \in A$ 並假設

$$f(a,b,c) = 0 \quad \text{且} \quad \frac{\partial f}{\partial z}(a,b,c) \neq 0.$$

那麼會存在 $\alpha, \beta, \gamma > 0$ 使得對於所有 $(x,y) \in (a-\alpha, a+\alpha) \times (b-\beta, b+\beta)$，方程式 $f(x,y,z) = 0$ 會在 $(c-\gamma, c+\gamma)$ 中有唯一的解 $z = \varphi(x,y)$。此外，函數 $\varphi$ 在 $(a-\alpha, a+\alpha) \times (b-\beta, b+\beta)$ 上會是 $\mathcal{C}^k$ 類的，且我們有

$$\begin{aligned}\frac{\partial \varphi}{\partial x}(x,y) &= -\frac{\partial f}{\partial x}(x,y,\varphi(x,y)) \Big/ \frac{\partial f}{\partial z}(x,y,\varphi(x,y)),\\ \frac{\partial \varphi}{\partial y}(x,y) &= -\frac{\partial f}{\partial y}(x,y,\varphi(x,y)) \Big/ \frac{\partial f}{\partial z}(x,y,\varphi(x,y)).\end{aligned}$$

**範例 4.4.15：** 我們考慮一個 $\mathcal{C}^\infty$ 函數 $f : \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto \sin(y) + xy^4 + x^2$。我們想要討論在 $(x,y) = (0,0)$ 附近，$f(x,y) = 0$ 的圖還有相關的漸進行為。

- 我們不難檢查 $f(0,0) = 0$。$f$ 的偏微分寫做

  $$\frac{\partial f}{\partial x}(x,y) = y^4 + 2x \quad \text{以及} \quad \frac{\partial f}{\partial y}(x,y) = \cos(y) + 4xy^3.$$

  我們有 $\frac{\partial f}{\partial x}(0,0) = 0$ 以及 $\frac{\partial f}{\partial y}(0,0) = 1$。因此，從系理 4.4.13 我們可以推得存在 $\alpha, \beta > 0$ 和 $\mathcal{C}^\infty$ 函數 $\varphi : (-\alpha, \alpha) \to \mathbb{R}$ 使得對於每個 $x \in (-\alpha, \alpha)$，$y = \varphi(x)$ 會是 $f(x,y) = 0$ 在

$(-\beta, \beta)$ 中唯一的解。

- 接著來計算 $\varphi$ 在 $0$ 附近的 Taylor 展開式。首先，從上面計算，我們得知 $\varphi(0) = 0$ 還有 $\varphi'(0) = 0$，所以當 $x \to 0$ 時，我們有 $\varphi(x) = \mathcal{O}(x^2)$。如果要得到更高階的展開，我們可以把展開式 $\sin(y) = y + \mathcal{O}(y^3)$ 當 $y = \varphi(x) \to 0$ 帶入到 $f(x, \varphi(x)) = 0$ 裡面。我們得到

$$\begin{aligned} \varphi(x) &= \varphi(x) - \sin(\varphi(x)) - x\varphi(x)^4 - x^2 \\ &= \mathcal{O}(\varphi(x)^3) - x^2 \\ &= -x^2 + \mathcal{O}(x^6) = -x^2(1 + \mathcal{O}(x^4)). \end{aligned}$$

如果我們想要得到 $\varphi$ 更高次的展開，我們需要把 $\sin$ 做更高階的展開，也就是 $\sin(y) = y - \frac{y^3}{6} + \mathcal{O}(y^5)$ 當 $y = \varphi(x) \to 0$。我們得到

$$\begin{aligned} \varphi(x) &= \varphi(x) - \sin(\varphi(x)) - x\varphi(x)^4 - x^2 \\ &= -x^2 - \frac{\varphi(x)^3}{6} + \mathcal{O}(\varphi(x)^5) - x\varphi(x)^4 \\ &= -x^2 + \frac{x^6}{6}(1 + \mathcal{O}(x^4)) + \mathcal{O}(x^{10}) - x^9(1 + \mathcal{O}(x^4)) \\ &= -x^2 + \frac{x^6}{6} - x^9 + \mathcal{O}(x^{10}). \end{aligned}$$

我們可以繼續這樣下去，藉由 $\sin$ 更高階的展開，得到 $\varphi$ 更高階的展開。

## 第四小節 條件極值

令 $A \subseteq \mathbb{R}^n$ 為開集且 $f : A \to \mathbb{R}$ 為函數。令 $g_1, \ldots, g_r : A \to \mathbb{R}$ 為函數以及

$$\Gamma = \{x \in A : g_1(x) = \cdots = g_r(x) = 0\}.$$

我們想要找出 $f$ 在 $\Gamma$ 上的極值。這樣的問題稱作條件極值。

**定理 4.4.16 ：** 假設 $f, g_1, \ldots, g_r$ 為 $\mathcal{C}^1$ 類函數。假設 $f_{|\Gamma}$ 在 $a \in \Gamma$ 有局部極值，且 $\mathrm{d}g_{1,a}, \ldots, \mathrm{d}g_{r,a}$ 是線性獨立的，那麼存在 $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ 使得

$$\mathrm{d}f_a = \lambda_1 \,\mathrm{d}g_{1,a} + \cdots + \lambda_r \,\mathrm{d}g_{r,a}. \tag{4.31}$$

**註解 4.4.17 ：** 我們把係數 $\lambda_i$ 稱作拉格朗日常數。這些係數有唯一性，因為線性泛函 $\mathrm{d}g_{1,a}, \ldots, \mathrm{d}g_{r,a}$ 是線性獨立的。

**證明：** 令 $s = n - r$ 並記 $\mathbb{R}^n = \mathbb{R}^s \times \mathbb{R}^r$。$\mathbb{R}^n$ 中的元素可以寫成 $(x, y) = (x_1, \ldots, x_s; y_1, \ldots, y_r)$。令 $a = (x_a, y_a) \in \mathbb{R}^n$，其中 $x_a \in \mathbb{R}^s$ 且 $y_a \in \mathbb{R}^r$。

首先，我們注意到由於 $(\mathrm{d}g_{i,a})_{1\leqslant i\leqslant r}$ 是線性獨立的，且線性泛函 $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ 的空間維度會等於 $n$，我們必然有 $r \leqslant n$。如果 $n = r$，那麼這個定理顯然成立，因為 $(\mathrm{d}g_{i,a})_{1\leqslant i\leqslant r}$ 構成基底。因此接下來，我們可以假設 $r \leqslant n - 1$，也就是 $s \geqslant 1$。

根據 $(\mathrm{d}g_{i,a})_{1\leqslant i\leqslant r}$ 的線性獨立性，$g = (g_1, \ldots, g_r)$ 在 $a$ 的 Jacobi 矩陣的秩會是 $r$。不失一般性，我們可以假設下面這個 $r \times r$ 子矩陣的行列式是非零的：

$$\det\left(\frac{\partial g_i}{\partial y_j}(a)\right)_{1\leqslant i,j\leqslant r} \neq 0.$$

因此，從系理 4.4.12 我們得知，存在 $\mathbb{R}^s$ 中包含 $x_a$ 的開集 $X$、$\mathbb{R}^n$ 中包含 $a = (x_a, y_a)$ 的開集 $W$，以及 $\mathcal{C}^1$ 類函數 $\varphi = (\varphi_1, \ldots, \varphi_r) : X \to \mathbb{R}^r$ 使得

$$g(x, y) = 0 \text{ 滿足 } x \in X \text{ 且 } (x, y) \in W \quad \Leftrightarrow \quad y = \varphi(x).$$

換句話說，對於 $x \in X$，$\Gamma = \{z : g(z) = 0\}$ 中的元素可以寫成 $(x, \varphi(x))$。令 $h(x) = f(x, \varphi(x))$，根據假設，他在 $x = a$ 有局部最大值。這給我們

$$0 = \frac{\partial h}{\partial x_i}(a) = \frac{\partial f}{\partial x_i}(a) + \sum_{j=1}^{r} \frac{\partial \varphi_j}{\partial x_i}(x_a)\frac{\partial f}{\partial y_j}(a), \quad \forall i = 1, \ldots, s. \tag{4.32}$$

此外，如果把式子 $g(x, \varphi(x)) = 0$ 微分，我們得到

$$0 = \frac{\partial g_k}{\partial x_i}(a) + \sum_{j=1}^{r} \frac{\partial \varphi_j}{\partial x_i}(x_a)\frac{\partial g_k}{\partial y_j}(a), \quad \forall k = 1, \ldots, r, \forall i = 1, \ldots, s. \tag{4.33}$$

把式 (4.32) 和式 (4.33) 寫成矩陣形式，我們得知矩陣

$$M = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) & \ldots & \frac{\partial f}{\partial x_s}(a) & \frac{\partial f}{\partial y_1}(a) & \ldots & \frac{\partial f}{\partial y_r}(a) \\ \frac{\partial g_1}{\partial x_1}(a) & \ldots & \frac{\partial g_1}{\partial x_s}(a) & \frac{\partial g_1}{\partial y_1}(a) & \ldots & \frac{\partial g_1}{\partial y_r}(a) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_r}{\partial x_1}(a) & \ldots & \frac{\partial g_r}{\partial x_s}(a) & \frac{\partial g_r}{\partial y_1}(a) & \ldots & \frac{\partial g_r}{\partial y_r}(a) \end{pmatrix}.$$

最前面 $s$ 行是最後 $r$ 行的線性組合，因此我們得到 $\operatorname{rank} M \leqslant r$。由於 $\operatorname{rank} M^t = \operatorname{rank} M$，這代表著 $M$ 中的 $r + 1$ 列是線性不獨立的，也就是說，會存在不會全部同時是零的 $\mu_0, \ldots, \mu_r$，使得

$$\mu_0 \,\mathrm{d}f_a + \mu_1 \,\mathrm{d}g_{1,a} + \cdots + \mu_r \,\mathrm{d}g_{r,a} = 0. \tag{4.34}$$

根據 $(\mathrm{d}g_{i,a})_{1\leqslant i\leqslant r}$ 是線性獨立的假設，我們得知 $\mu_0 \neq 0$，所以我們可以把式 (4.34) 同除 $\mu_0$，此定理得證。 □

**範例 4.4.18**：求函數 $f(x,y) = x^2 - 2x + 4y^2 + 8y$ 在第一象限中，也就是 $(x,y) \in A := \mathbb{R}_{\geqslant 0} \times \mathbb{R}_{\geqslant 0}$，以及條件 $g(x,y)=0$，其中 $g(x,y) = x + 2y - 7 = 0$，之下的最小值與最大值。由於定義域 $A$ 在 $\mathbb{R}^2$ 不是閉集，我們需要區分內點以及其他點。$f$ 的極值會在滿足式 (4.31) 的點 $a \in \mathring{A} = \mathbb{R}_{>0} \times \mathbb{R}_{>0}$，或是在定理 4.4.16 中沒有討論到的點 $A \backslash \mathring{A} = (\{0\} \times \mathbb{R}_{\geqslant 0}) \cup (\mathbb{R}_{\geqslant 0} \times \{0\})$ 上碰到。

- 我們先找內點 $(x,y) \in \mathring{A}$，使得他滿足 $g(x,y)=0$ 以及方程式 $\mathrm{d}f_{(x,y)} = \lambda\, \mathrm{d}g_{(x,y)}$ 會有非零的解 $\lambda \in \mathbb{R}$。首先，我們要解

$$\begin{cases} \frac{\partial f}{\partial x}(x,y) &= \lambda \frac{\partial g}{\partial x}(x,y), \\ \frac{\partial f}{\partial y}(x,y) &= \lambda \frac{\partial g}{\partial y}(x,y), \end{cases} \quad \Leftrightarrow \quad \begin{cases} 2x-2 &= \lambda, \\ 8y+8 &= 2\lambda. \end{cases}$$

  因此，我們得到 $x = 2y+3$。我們把這個帶回條件 $g(x,y)=0$，得到 $(x,y)=(5,1)$。我們計算在這個點的值 $f(5,1)=27$。

- 在 $A \backslash \mathring{A}$ 中滿足 $g(x,y)=0$ 的點 $(x,y)$ 會是 $(x,y)=(7,0)$ 或 $(0,\frac{7}{2})$。我們計算函數在這兩點的值：$f(7,0)=35$ 以及 $f(0,\frac{7}{2})=77$。

從上述計算，我們得知在 j 第一象限中，以及條件 $g(x,y)=0$ 之下，$f$ 在 $(0,\frac{7}{2})$ 有最大值，取值為 77，他在 $(5,1)$ 有最小值，取值為 27。