

附录 A Hilbert 空间

考虑空间 H ，其元素为实随机函数 $X_1(t, \omega), X_2(t, \omega), \dots, \omega \in \Omega$ ，其中 Ω 为样本空间。假定在这一空间中，所有的元素 $X_i(t) \stackrel{\text{def}}{=} X_i(t, \omega)$ ， $i = 1, 2, \dots$ 都具有有限的二阶矩，即

$$E\{X_i^2(t)\} < \infty, \quad i = 1, 2, \dots \quad (\text{附A.1})$$

线性空间 H 定义为：若 $X_i(t) \in H$ 和 $X_j(t) \in H$ ，则 $\alpha X_i(t) + \beta X_j(t)$ 也一定属于 H 空间。下面约定 H 为线性空间。线性空间 H 的内积和范数定义为

$$\langle X_i, X_j \rangle \stackrel{\text{def}}{=} E\{X_i(t)X_j(t)\} \quad (\text{附A.2})$$

$$\|X_i(t)\|^2 \stackrel{\text{def}}{=} E\{X_i^2(t)\} \quad (\text{附A.3})$$

假定在 H 空间存在以下基本公理：

A1: $X_i(t) + X_j(t) = X_j(t) + X_i(t)$ (交换律)

A2: $[X_i(t) + X_j(t)] + X_k(t) = X_i(t) + [X_j(t) + X_k(t)]$ (结合律)

A3: $(\alpha + \beta)X_i(t) = \alpha X_i(t) + \beta X_i(t)$ (分配律)

A4: $\alpha[X_i(t) + X_j(t)] = \alpha X_i(t) + \alpha X_j(t)$

A5: 存在一个零元素 $\theta \in H$ 使得

$$X_i(t) + \theta = X_i(t) \quad \forall X_i \in H$$

A6: $0X_i(t) = 0$ 和 $1X_i(t) = X_i(t)$ 。

下面是关于范数的公理。

B1: 范数是非负的，即 $\|X_i(t)\| \geq 0$ ， $\forall X_i(t) \in H$ ，且若 $\|X_i(t)\| = 0$ ，则 $X_i(t) = 0$ 。

B2: 范数满足三角不等式，即 $\|X_i(t) + X_j(t)\| \leq \|X_i(t)\| + \|X_j(t)\|$ 对所有 $X_i(t), X_j(t) \in H$ 成立。

B3: $\|\alpha X_i(t)\| = |\alpha| \cdot \|X_i(t)\|$ 对所有标量 α 和 $x_i(t) \in H$ 成立。

以下是内积服从的公理。

C1: $\langle X_i(t), X_j(t) \rangle = \langle X_j(t), X_i(t) \rangle$

C2: $\langle X_i + X_j(t), X_k(t) \rangle = \langle X_i(t), X_k(t) \rangle + \langle X_j(t), X_k(t) \rangle$

$$C3: \lambda \langle X_i(t), X_j(t) \rangle = \langle \lambda X_i(t), X_j(t) \rangle$$

C4: 内积是非负的, 即 $\langle X_i(t), X_j(t) \rangle \geq 0$, 并且若 $\langle X_i(t), X_j(t) \rangle = 0$, 则 $X_i(t) = 0$.

若 H 满足公理 A1~A6, 则称 H 为线性向量空间. 若 H 满足公理 A1~A6 和 B1~B3, 则称 H 为赋范线性空间.

若 $\lim_{n \rightarrow \infty} X_n(t) \rightarrow X(t)$ 在均方意义下成立, 且 $X(t) \in H$, 则称空间 H 是完备空间. 一个完备的赋范线性空间 H 称为 Banach 空间. 如果完备的 Banach 空间进一步满足公理 C1~C4, 则称空间 H 为 Hilbert 空间. 简言之, 一个完备的线性空间 H 若满足公理 A1~A6、公理 B1~B3 和公理 C1~C4, 则 H 是 Hilbert 空间.

考虑每个元素均为向量的 Hilbert 空间. 记 $x_i(t) = [X_{i1}(t), X_{i2}(t), \dots, X_{in}(t)]^T$, 并且线性向量空间 H 的每一个向量都具有 n 个元素. 线性向量空间 H 的内积和范数分别定义为

$$\langle x_i(t), x_j(t) \rangle \stackrel{\text{def}}{=} E\{x_i^T(t)x_j(t)\} \quad (\text{附A.4})$$

$$\|x_i(t)\|^2 \stackrel{\text{def}}{=} E\{x_i^T(t)x_i(t)\} \geq 0 \quad (\text{附A.5})$$

两个向量 $x_i(t)$ 和 $x_j(t)$ 正交, 若它们的内积等于零, 即

$$\langle x_i(t), x_j(t) \rangle = E\{x_i^T(t)x_j(t)\} = 0, \quad i \neq j \quad (\text{附A.6})$$

并记作 $x_i(t) \perp x_j(t)$, $i \neq j$.

为简便计, 引入符号 $x_i(t) = x$ 和 $y_i(t) = y$.

命题 1 若 $x \perp y$, 则 $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

证明 由公理 A2 和 C2 知

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \quad (\text{附A.7})$$

由于 x 和 y 正交, 所以 $\langle x, y \rangle = E\{x^T y\} = 0$. 又由公理 C1 知 $\langle y, x \rangle = \langle x, y \rangle = 0$. 将这一结果代入式 (附 A.7) 立即得

$$\|x + y\|^2 = \langle x, x \rangle + \langle y, y \rangle = \|x\|^2 + \|y\|^2 \quad (\text{附A.8})$$

这就完成了本命题的证明. ■

这一命题也称 Pythagorean 定理.

命题 2 $|\langle x, y \rangle|^2 \leq \|x\| \|y\|$, 等号成立当且仅当 $x = \lambda y$ 或 $y = 0$ (零向量).

证明 由公理 B1 知

$$\langle x - \lambda y, x - \lambda y \rangle \geq 0 \quad (\text{附A.9})$$

式中 $\lambda \neq 0$ 为一标量。注意到

$$\langle x - \lambda y, x - \lambda y \rangle = \langle x, x \rangle - \lambda \langle y, x \rangle - \lambda \langle x, y \rangle + \lambda^2 \langle y, y \rangle$$

利用公理 B1~B3 和 C1~C4, 并选择

$$\lambda = \frac{\langle x, y \rangle}{\|y\|^2}$$

则由式 (附 A.9) 知

$$\begin{aligned} \langle x - \lambda y, x - \lambda y \rangle &= \langle x, x \rangle - |\langle x, y \rangle|^2 / \|y\|^2 \\ &= \langle x, x \rangle - |\langle x, y \rangle|^2 / \langle y, y \rangle \geq 0 \end{aligned}$$

由此得 $|\langle x, y \rangle|^2 \leq \|x\| \|y\|$ 。显然, 当 $x = \lambda y$ 或 $y = 0$ 时, 不等式取等号。 ■

不等式 $|\langle x, y \rangle|^2 \leq \|x\| \|y\|$ 称为 Cauchy-Schwartz 不等式。

下面的命题称为平行四边形法则。

命题 3 $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$ 。

证明 具体计算范数, 得

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\ \|x - y\|^2 &= \langle x - y, x - y \rangle = \langle x, x \rangle - 2\langle x, y \rangle + \langle y, y \rangle \end{aligned}$$

两式相加, 即有

$$\|x + y\|^2 + \|x - y\|^2 = 2\langle x, x \rangle + 2\langle y, y \rangle = 2(\|x\|^2 + \|y\|^2) \quad (\text{附A.10})$$

这即是希望证明的结果。 ■

定义 1 称 x_n 收敛为 x 若 $\|x_n - x\| \rightarrow 0$ 当 $n \rightarrow \infty$ 时。

下面的命题表明, 范数具有连续性。

命题 4 若 $x_n \rightarrow x$ 和 $y_n \rightarrow y$, 则 $\langle x_n, y_n \rangle \rightarrow \langle x, y \rangle$ 。

证明 若 $x_n \rightarrow x$ 和 $y_n \rightarrow y$, 则 x_n 和 y_n 必然是有界的向量, 即

$$\|x_n\| \leq m_1 \quad \text{和} \quad \|y_n\| \leq m_2$$

且

$$\|x_n - x\| \rightarrow 0, \quad \text{当 } n \rightarrow \infty$$

$$\|y_n - y\| \rightarrow 0, \quad \text{当 } n \rightarrow \infty$$

考查

$$\begin{aligned}\langle x, y \rangle &= \langle x - x_n + x_n, y - y_n + y_n \rangle \\ &= \langle x - x_n, y - y_n \rangle + \langle x - x_n, y_n \rangle + \langle x_n, y - y_n \rangle + \langle x_n, y_n \rangle\end{aligned}\quad (\text{附A.11})$$

然而, 由命题 2 及题意知, 上该式右边前三项分别为

$$\left. \begin{aligned}|\langle x - x_n, y - y_n \rangle|^2 &\leq \|x - x_n\| \cdot \|y - y_n\| \rightarrow 0 \quad \text{当 } n \rightarrow \infty \\ \langle x - x_n, y_n \rangle &\leq \|y_n\| \cdot \|x - x_n\| \leq m_2 \|x - x_n\| \rightarrow 0 \quad \text{当 } n \rightarrow \infty \\ \langle x_n, y - y_n \rangle &\leq \|x_n\| \cdot \|y - y_n\| \leq m_1 \|y - y_n\| \rightarrow 0 \quad \text{当 } n \rightarrow \infty\end{aligned}\right\} \quad (\text{附A.12})$$

将以上公式代入式 (附 A.11) 可知, 下面的收敛结果成立:

$$\langle x_n, y_n \rangle \rightarrow \langle x, y \rangle \quad \text{当 } n \rightarrow \infty \quad (\text{附A.13})$$

这就完成了本命题的证明。■

定义 2 H 空间的线性流形 $L \subset H$ 是 H 空间内满足下列条件的向量子集合: 若 $x_i \in L$ 和 $x_j \in L$, 则 $\alpha_1 x_i + \alpha_2 x_j \in L$ 对于任何一对标量 α_1 和 α_2 恒成立。

定义 3 一向量组 $\{x_i, i = 1, \dots, n\}$ 称作线性独立, 若对已知的常数 $c_1, \dots, c_n$, 等式

$$\sum_{i=1}^n c_i x_i = 0$$

意味着所有 $c_1, \dots, c_n$ 皆等于零。

定义 4 若 H 空间内的线性独立的向量产生或张成 H , 则这些向量称为 H 空间的基 (向量)。特别地, 若基向量 $\{x_i\}$ 满足正交条件

$$\langle x_i, x_j \rangle = 0, \quad \forall i \neq j \quad (\text{附A.14})$$

则称这些基向量为正交基向量。进一步地, 若正交基向量的范数为 1, 即

$$\|x_i\| = 1 \quad \forall i \quad (\text{附A.15})$$

则称这样的基向量为标准正交基向量。

附录 B Cauchy-Schwartz 不等式

对任意复函数 $f(x)$ 和 $g(x)$, Cauchy-Schwartz 不等式

$$n \left| \int_a^b f(x)g(x)dx \right|^2 \leq \int_a^b |f(x)|^2 dx \cdot \int_a^b |g(x)|^2 dx \quad (\text{附B.1})$$

恒成立, 并且当且仅当在积分区间 $[a, b]$ 内, $f(x)$ 与 $g^*(x)$ 成正比, 即

$$f(x) = cg^*(x) \quad (\text{附B.2})$$

时, 式 (附 B.1) 中的等号才成立。

证明 实现, 证明式 (附 B.1) 对实函数成立。为此, 构造二次式

$$I(y) = \int_a^b [f(x) - yg(x)]^2 dx \quad (\text{附B.3})$$

$$= \int_a^b f^2(x)dx - 2y \int_a^b f(x)g(x)dx + y^2 \int_a^b g^2(x)dx \quad (\text{附B.4})$$

回顾二次式 $I(y) = A - 2By + cy^2$ 的性质: 若 $I(y) \geq 0, \forall y$, 则 $B^2 - AC \leq 0$ 。

由式 (附 B.3) 可以看出 $I(y) \geq 0$, 因此

$$\begin{aligned} B^2 - AC &= \left| \int_a^b f(x)g(x)dx \right|^2 - \int_a^b f^2(x)dx \cdot \int_a^b g^2(x)dx \\ &\leq 0 \end{aligned}$$

或

$$\left| \int_a^b f(x)g(x)dx \right|^2 \leq \int_a^b f^2(x)dx \cdot \int_a^b g^2(x)dx$$

另一方面, 若 $B^2 - AC = 0$ 即式 (附 B.1) 取等号, 则二次式 $I(y)$ 必定有一个实的重根。这就是说, 对某个 $y = c$, 有

$$I(y) = \int_a^b [f(x) - cg(x)]^2 dx = 0$$

即当不等式 (附 B.1) 取等号时, 必有 $f(x) = cg(x)$ 。因此, Cauchy-Schwartz 不等式对实函数成立。

当 $f(x)$ 和 $g(x)$ 为复函数时, 令 $\int_a^b |f(x)|^2 dx = A$, $\int_a^b f(x)g(x)dx = B$ 和 $\int_a^b |g(x)|^2 dx = C$, 并令 y 是任意实数. 于是, 有

$$\begin{aligned} & \int_a^b |f(x) - yg^*(x)|^2 dx \\ &= \int_a^b |f(x)|^2 dx - 2y \int_a^b |f(x)g(x)| dx + y^2 \int_a^b |g(x)|^2 dx \\ &= A - 2By + Cy^2 \end{aligned}$$

这是一个非负的二次多项式, 且 A, B, C 均为实数. 由前面已证明的结果知, Cauchy-Schwartz 不等式 (附 B.1) 对复函数 $f(x)$ 和 $g(x)$ 也成立.

附录 C 相对于复向量的导数

在信号处理中，代价函数常定义为误差的均方值或平方和等形式，为一实的标量；而待估计的参数向量或滤波器权向量往往又是复向量。代价函数 J 的优化（即最小化）就要求标量 J 相对于复（列）向量 $\mathbf{w}$ 的偏导。

令 $\mathbf{w} = [w_1, \dots, w_n]^T$ 是一 n 维复列向量，标量 J 相对于列向量 $\mathbf{w}$ 和 $\mathbf{w}^*$ 的偏导分别定义为

$$\frac{\partial J}{\partial \mathbf{w}} = \left[\frac{\partial J}{\partial w_1}, \dots, \frac{\partial J}{\partial w_n} \right]^T \quad (\text{附C.1})$$

和

$$\frac{\partial J}{\partial \mathbf{w}^*} = \left[\frac{\partial J}{\partial w_1^*}, \dots, \frac{\partial J}{\partial w_n^*} \right]^T \quad (\text{附C.2})$$

偏导 $\partial J / \partial \mathbf{w}$ 称为标量 J 相对于向量 $\mathbf{w}$ 的偏导，而 $\partial J / \partial \mathbf{w}^*$ 称为 J 相对于向量 $\mathbf{w}$ 的共轭偏导。

令

$$w_k = x_k + jy_k, \quad k = 0, 1, \dots, M-1$$

即 w_k 的实部为 x_k 、虚部为 y_k ，复数的形式导数定义为

$$\frac{\partial}{\partial w_k} = \frac{1}{2} \left(\frac{\partial}{\partial x_k} - j \frac{\partial}{\partial y_k} \right) \quad (\text{附C.3})$$

和

$$\frac{\partial}{\partial w_k^*} = \frac{1}{2} \left(\frac{\partial}{\partial x_k} + j \frac{\partial}{\partial y_k} \right) \quad (\text{附C.4})$$

这一定义满足以下两个性质：

$$\frac{\partial w_k}{\partial w_k} = 1 \quad (\text{附C.5})$$

$$\frac{\partial w_k}{\partial w_k^*} = \frac{\partial w_k^*}{\partial w_k} = 0 \quad (\text{附C.6})$$

利用式 (附 C.3) 和式 (附 C.4), 即可写出复向量的形式导数

$$\frac{\partial}{\partial \mathbf{w}} = \frac{1}{2} \begin{bmatrix} \frac{\partial}{\partial x_1} - j \frac{\partial}{\partial y_1} \\ \vdots \\ \frac{\partial}{\partial x_n} - j \frac{\partial}{\partial y_n} \end{bmatrix} \quad (\text{附C.7})$$

和

$$\frac{\partial}{\partial \mathbf{w}^*} = \frac{1}{2} \begin{bmatrix} \frac{\partial}{\partial x_1} + j \frac{\partial}{\partial y_1} \\ \vdots \\ \frac{\partial}{\partial x_n} + j \frac{\partial}{\partial y_n} \end{bmatrix} \quad (\text{附C.8})$$

若 $\mathbf{J} = [J_1, \dots, J_m]$ 为一行向量 (实或复向量), 则 $\mathbf{J}$ 相对于复列向量 $\mathbf{w}$ 的偏导为一 $n \times m$ 矩阵, 并定义为

$$\frac{\partial \mathbf{J}}{\partial \mathbf{w}} = \begin{bmatrix} \frac{\partial J_1}{\partial w_1} & \dots & \frac{\partial J_m}{\partial w_1} \\ \vdots & \vdots & \vdots \\ \frac{\partial J_1}{\partial w_n} & \dots & \frac{\partial J_m}{\partial w_n} \end{bmatrix} \quad (\text{附C.9})$$

类似地, $\mathbf{J}$ 相对于复列向量 $\mathbf{w}^*$ 的偏导定义为

$$\frac{\partial \mathbf{J}}{\partial \mathbf{w}^*} = \begin{bmatrix} \frac{\partial J_1}{\partial w_1^*} & \dots & \frac{\partial J_m}{\partial w_1^*} \\ \vdots & \vdots & \vdots \\ \frac{\partial J_1}{\partial w_n^*} & \dots & \frac{\partial J_m}{\partial w_n^*} \end{bmatrix} \quad (\text{附C.10})$$

虽然原理上也可定义 $m \times 1$ 列向量对 $n \times 1$ 列向量的偏导, 其结果为 $mn \times 1$ 列向量, 但在信号处理中不涉及这类偏导。

根据定义式 (附 C.9)、式 (附 C.10) 以及式 (附 C.5)、式 (附 C.6), 不难验证

$$\frac{\partial \mathbf{w}^T}{\partial \mathbf{w}} = \frac{\partial \mathbf{w}^H}{\partial \mathbf{w}^*} = \mathbf{I} \quad (\text{附C.11})$$

$$\frac{\partial \mathbf{w}^T}{\partial \mathbf{w}^*} = \frac{\partial \mathbf{w}^H}{\partial \mathbf{w}} = \mathbf{O} \quad (\text{附C.12})$$

式中 $\mathbf{I}$ 和 $\mathbf{O}$ 分别为单位矩阵和零矩阵。式 (附 C.12) 表明, 复向量 $\mathbf{w}$ 相对于其共轭向量 $\mathbf{w}^*$ 可视为一常数, 而 $\mathbf{w}^*$ 相对于 $\mathbf{w}$ 也可视为一常数。

例 1 令 $J_1 = \mathbf{p}^H \mathbf{w}$ 为一实函数, 其中 $\mathbf{p}$ 和 $\mathbf{w}$ 均为 $n \times 1$ 复向量。于是, J_1 相对于 $\mathbf{w}^*$ 和 $\mathbf{w}$ 的偏导分别为

$$\frac{\partial J_1}{\partial \mathbf{w}^*} = \frac{\partial}{\partial \mathbf{w}^*} (\mathbf{p}^H \mathbf{w}) = \frac{\partial \mathbf{w}^T}{\partial \mathbf{w}^*} \mathbf{p}^* = \mathbf{O} \cdot \mathbf{p}^* = 0 \quad (\text{附C.13})$$

$$\frac{\partial J_1}{\partial \mathbf{w}} = \frac{\partial}{\partial \mathbf{w}} (\mathbf{p}^H \mathbf{w}) = \frac{\partial \mathbf{w}^T}{\partial \mathbf{w}} \mathbf{p}^* = \mathbf{I} \cdot \mathbf{p}^* = \mathbf{p}^* \quad (\text{附C.14})$$

例 2 若 $\mathbf{p}$ 和 $\mathbf{w}$ 同为 $n \times 1$ 复向量, 并且 $J_2 = \mathbf{w}^H \mathbf{p}$ 为一实函数, 则

$$\frac{\partial J_2}{\partial \mathbf{w}^*} = \frac{\partial}{\partial \mathbf{w}^*} (\mathbf{w}^H \mathbf{p}) = \frac{\partial \mathbf{w}^H}{\partial \mathbf{w}^*} \mathbf{p} = \mathbf{I} \cdot \mathbf{p} = \mathbf{p} \quad (\text{附C.15})$$

$$\frac{\partial J_2}{\partial \mathbf{w}} = \frac{\partial}{\partial \mathbf{w}} (\mathbf{w}^H \mathbf{p}) = \frac{\partial \mathbf{w}^H}{\partial \mathbf{w}} \mathbf{p} = \mathbf{O} \cdot \mathbf{p} = 0 \quad (\text{附C.16})$$

以上两个例子说明, 究竟是选择对 $\mathbf{w}$ 还是 $\mathbf{w}^*$ 求偏导, 完全决定于实函数本身的形式。

例 3 考虑二次型代价函数

$$J(\mathbf{w}) = \mathbf{w}^H \mathbf{R} \mathbf{w} \quad (\text{附C.17})$$

式中 $\mathbf{R}$ 为一共轭对称的 Hermitian 矩阵。此时, 我们有

$$\frac{\partial J}{\partial \mathbf{w}^*} = \frac{\partial}{\partial \mathbf{w}^*} (\mathbf{w}^H \mathbf{R} \mathbf{w}) = \frac{\partial \mathbf{w}^H}{\partial \mathbf{w}^*} \mathbf{R} \mathbf{w} = \mathbf{R} \mathbf{w} \quad (\text{附C.18})$$

注意, 如果 $\mathbf{R}$ 是实对称矩阵, 而 $\mathbf{w}$ 也是一实向量时, 则

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{w}} &= \frac{\partial}{\partial \mathbf{w}} (\mathbf{w}^T \mathbf{R} \mathbf{w}) = \frac{\partial \mathbf{w}^T}{\partial \mathbf{w}} \mathbf{R} \mathbf{w} + \frac{\partial (\mathbf{R} \mathbf{w})^T}{\partial \mathbf{w}} \mathbf{w} \\ &= \mathbf{I} \cdot \mathbf{R} \mathbf{w} + \mathbf{R} \mathbf{w} = 2\mathbf{R} \mathbf{w} \end{aligned} \quad (\text{附C.19})$$

考虑实代价函数 $J(\mathbf{w})$, 其中 $\mathbf{w} = [w_0, w_1, \dots, w_{M-1}]^T$, 则 $J(\mathbf{w})$ 的梯度定义为

$$\nabla J = \begin{bmatrix} \frac{\partial J}{\partial x_0} + j \frac{\partial J}{\partial y_0} \\ \frac{\partial J}{\partial x_1} + j \frac{\partial J}{\partial y_1} \\ \vdots \\ \frac{\partial J}{\partial x_{M-1}} + j \frac{\partial J}{\partial y_{M-1}} \end{bmatrix} \quad (\text{附C.20})$$

将定义式 (附 C.20) 与 (附 C.8) 作一比较知, 共轭偏导 $\partial J / \partial \mathbf{w}^*$ 与梯度向量 ∇J 存在以下关系:

$$\nabla J = 2 \frac{\partial J}{\partial \mathbf{w}^*} \quad (\text{附C.21})$$

因此,除了相差一个尺度因子 2 以外,梯度向量与其轭偏导相同。

例 4 考查超定矩阵方程 $\mathbf{A}\mathbf{w} = \mathbf{b}$ 的求解。引入对数似然函数

$$l(\hat{\mathbf{w}}) = C - \frac{1}{\sigma^2} \mathbf{e}^H \mathbf{e} \quad (\text{附C.22})$$

式中 C 为一常数,而 $\mathbf{e}$ 为误差向量,定义为

$$\mathbf{e} = \mathbf{b} - \mathbf{A}\hat{\mathbf{w}} \quad (\text{附C.23})$$

将式 (附 C.23) 代入式 (附 C.22) 得

$$l(\hat{\mathbf{w}}) = C - \frac{1}{\sigma^2} \mathbf{b}^H \mathbf{b} + \frac{1}{\sigma^2} \mathbf{b}^H \mathbf{A} \hat{\mathbf{w}} + \frac{1}{\sigma^2} \hat{\mathbf{w}}^H \mathbf{A}^H \mathbf{b} - \frac{1}{\sigma^2} \hat{\mathbf{w}}^H \mathbf{A}^H \mathbf{A} \hat{\mathbf{w}} \quad (\text{附C.24})$$

求对数似然函数相对于 $\mathbf{w}$ 的共轭偏导,则

$$\frac{\partial l}{\partial \hat{\mathbf{w}}^*} = \frac{1}{\sigma^2} \mathbf{A}^H \mathbf{b} - \frac{1}{\sigma^2} \mathbf{A}^H \mathbf{A} \hat{\mathbf{w}}$$

令其等于零,得

$$\mathbf{A}^H \mathbf{b} - \mathbf{A}^H \mathbf{A} \mathbf{w}_{\text{opt}} = 0 \quad \text{或} \quad \mathbf{A}^H \mathbf{A} \mathbf{w}_{\text{opt}} = \mathbf{A}^H \mathbf{b}$$

式中 $\mathbf{w}_{\text{opt}}$ 是使对数似然函数 $l(\hat{\mathbf{w}})$ 最大化的 $\hat{\mathbf{w}}$ 值。于是,若 $\mathbf{A}^H \mathbf{A}$ 非奇异,则

$$\mathbf{w}_{\text{opt}} = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{b} \quad (\text{附C.25})$$

这就是矩阵方程 $\mathbf{A}\mathbf{w} = \mathbf{b}$ 的最大似然解。