第三章 自适应波束形成

§3.1 引言

- **最优波束形成器：**

  假设已知信号的时间信息（参考信号）或空间信息（到达方向）

  假设已知二阶统计量R_x、R_n、R_s

- **自适应波束形成器：**

  假设信号的时间或空间信息已知(或可以通过训练序列得到准确的估计)

  二阶统计量未知，需要通过从训练序列中“学习”得到统计信息

- **盲自适应波束形成器：**

  信号的时间或空间信息未知 —— 需要从接收数据中估计出来

二阶统计量未知　　　　　——　也需要从接收数据中估计出来

自适应波束形成器可分为：

批处理算法——或分段数据处理器，从一段数据中估计相关矩阵$R_{\mathbf{x}}$或$R_{\mathbf{n}}$，并将估计值应用于最优波束形成器的公式中。被称为采样矩阵逆技术（SMI: Sample Matrix Inversion），或直接矩阵逆（DMI: Direct Matrix Inversion）技术。

递推算法——用迭代的方式实现矩阵的逆，被称为递推最小二乘算法（RLS：Recursive Least Square）。

基于梯度的算法——如最速下降法（SD: Steepest Descend 确定性梯度算法）和最小均方算法（LMS: Least Mean Square，随机梯度算法）。

自适应波束形成器的性能：

瞬态性能——分析算法的收敛速度

稳态性能——分析自适应算法达到稳态时的性能与相应最优波束形成器的差异

§3.2 空间谱矩阵估计

一、极大似然估计

对于由N个阵元组成的传感器阵，考虑K个频域快拍$\mathbf{X}_1$，$\mathbf{X}_2$，…，$\mathbf{X}_K$，其中$\mathbf{X}_k$为阵列在k时刻得到的N维接收信号矢量。可以将这些频域快拍表示成一个$N\times K$维的数据矩阵：

$$\tilde{\mathbf{X}} = \frac{1}{\sqrt{K}}\begin{bmatrix}\mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_K\end{bmatrix}$$

假设这些频域快拍是统计独立同分布的复高斯随机向量，则其联合概率密度为：

$$p_{\mathbf{X}}(\mathbf{X}_1,\mathbf{X}_2,\cdots,\mathbf{X}_K) = \prod_{k=1}^{K}\frac{\exp\left[-\mathbf{X}_k^H S_{\mathbf{x}}^{-1}\mathbf{X}_k\right]}{\pi\cdot\left|S_{\mathbf{x}}\right|}$$

取对数可得：

$$L_{\mathbf{X}} = -K\ln\pi - K\ln\left|S_{\mathbf{x}}\right| - \sum_{k=1}^{K}\mathbf{X}_k^H S_{\mathbf{x}}^{-1}\mathbf{X}_k$$

去掉常数项并经过简单运算得对数似然比为：

$$L\left(S_{\mathbf{x}}^{-1}\right)=-\ln\left|S_{\mathbf{x}}\right|-\frac{1}{K}\sum_{k=1}^{K}\mathbf{X}_k^H S_{\mathbf{x}}^{-1}\mathbf{X}_k=-\ln\left|S_{\mathbf{x}}\right|-tr\left(\frac{1}{K}\sum_{k=1}^{K}\mathbf{X}_k^H S_{\mathbf{x}}^{-1}\mathbf{X}_k\right)=-\ln\left|S_{\mathbf{x}}\right|-tr\left(S_{\mathbf{x}}^{-1}\cdot\frac{1}{K}\sum_{k=1}^{K}\mathbf{X}_k\mathbf{X}_k^H\right)$$

定义采样谱矩阵为：

$$C_{\mathbf{x}}=\frac{1}{K}\sum_{k=1}^{K}\mathbf{X}_k\mathbf{X}_k^H=\tilde{\mathbf{X}}\cdot\tilde{\mathbf{X}}^H$$

将采样谱矩阵代入频域快拍的对数似然比中：

$$L\left(S_{\mathbf{x}}^{-1}\right)=-\ln\left|S_{\mathbf{x}}\right|-tr\left[S_{\mathbf{x}}^{-1}C_{\mathbf{x}}\right]$$

令$\nabla L\left(S_{\mathbf{x}}^{-1}\right)/\nabla S_{\mathbf{x}}=0$，根据：

$$\frac{\partial}{\partial X}tr\left(X^{-1}A\right)=\left(-X^{-1}AX^{-1}\right)^T \quad 及 \quad \frac{\partial}{\partial X}\ln\left|X\right|=\left(X^{-1}\right)^T$$

可得空间谱矩阵的极大似然估计为：

$$\hat{S}_{\mathbf{x}}=C_{\mathbf{x}}$$

可见谱矩阵的极大似然估计就是频域快拍的采样相关矩阵。上述估计对空间谱矩阵的结构没有任何限制。$\hat{S}_{\mathbf{x}}$是 Hermitian 矩阵，即$\hat{S}_{\mathbf{x}}=\hat{S}_{\mathbf{x}}^H$。只要$K\geq N$，则$\hat{S}_{\mathbf{x}}$是正定矩阵。

二、结构化空间谱矩阵估计

如果对信号、干扰或噪声或阵列结构知道一些先验信息，则可以提高空间谱估计的性能——提高估计精度或提高收敛速度。

若阵列输入包括D个白高斯噪声中的不相关平面波信号，则空间谱矩阵具有如下的结构：

$$S_{\mathbf{x}} = V(\psi)\cdot S_{\mathbf{f}}\cdot V^{H}(\psi) + \sigma_w^2 \cdot I$$

其中$V(\psi) = [\mathbf{v}(\psi_1) | \ \cdots \ | \ \mathbf{v}(\psi_D)]$为$N\times D$维阵簇矩阵，$S_{\mathbf{f}} = diag\{\sigma_1^2, \sigma_2^2, \cdots, \sigma_D^2\}$。

对下面$2D+1$个参数进行极大似然估计(见[3]第 8 章)：

$$\boldsymbol{\theta} = [\psi_1, \ \cdots \ \psi_D, \ \sigma_1^2, \ \cdots \ \sigma_D^2, \ \sigma_w^2]$$

则可以利用参数估计方法得到空间谱矩阵的估计为：

$$\hat{S}_{\mathbf{x}} = V(\hat{\psi})\cdot \hat{S}_{\mathbf{f}}\cdot V^{H}(\hat{\psi}) + \hat{\sigma}_w^2 \cdot I$$

若阵列是均匀线阵，则根据[3]的 2.3 节可知，此时阵簇矢量为共轭对称的，即：

$$\mathbf{v}(\psi)=J\cdot\mathbf{v}^{H}(\psi)$$

其中 J 为交换矩阵。

这时，若频域快拍是统计独立同分布的复高斯随机向量，则用同样的方法可以证明空间谱矩阵的极大似然估计为：

$$\hat{S}_{\mathbf{x}}\triangleq\hat{S}_{\mathbf{x},fb}=C_{\mathbf{x},fb}=\frac{1}{2}\left(\tilde{\mathbf{X}}\cdot\tilde{\mathbf{X}}^{H}+J\cdot\tilde{\mathbf{X}}^{*}\cdot\tilde{\mathbf{X}}^{T}\cdot J\right)=\tilde{\mathbf{X}}_{fb}\cdot\tilde{\mathbf{X}}_{fb}^{H}$$

其中 $C_{\mathbf{x},fb}=\frac{1}{2}\left(C_{\mathbf{x}}+J\cdot C_{\mathbf{x}}^{*}\cdot J\right)$为对采样谱矩阵进行前后平均的结果，$\tilde{\mathbf{X}}_{fb}=\frac{1}{\sqrt{2}}\left[\tilde{\mathbf{X}}\quad|\quad J\cdot\tilde{\mathbf{X}}^{*}\right]$为 $N\times 2K$ 维数据矩阵。

$\hat{S}_{\mathbf{x},fb}$具有一些特殊的性质使得经过变换处理后的谱矩阵为实数矩阵，而对实数进行处理会节省 75%的计算量。

上述阵簇矢量的共轭对称性质对标准矩形平面阵、一些标准六角阵、阵元数为偶数的均匀园阵及均匀柱阵都成立。对于线阵，只要阵列单元相对于原点对称，并不要求均匀线阵。

对于**均匀线阵**，空间谱矩阵也是 Toeplitz 阵，该矩阵的所有对角元素都相同，$a_{ij}=a_{i-j}$。

例如，4×4 维 Toeplitz 矩阵为：

$$A=\begin{bmatrix} a_0 & a_{-1} & a_{-2} & a_{-3} & a_{-4} \\ a_1 & a_0 & a_{-1} & a_{-2} & a_{-3} \\ a_2 & a_1 & a_0 & a_{-1} & a_{-2} \\ a_3 & a_2 & a_1 & a_0 & a_{-1} \end{bmatrix}$$

利用 Toeplitz 阵的特点，人们也提出各种提高谱矩阵估计性能的方法，如 M. Miller 和 D. Snyder 于 1988 年提出的 EM(Expectation Maximization)方法。

上述对频域快拍求空间谱估计的各种方法同样可以用于对时域快拍$\mathbf{x}_k$的协方差矩阵进行估计：

$$\hat{R}_{\mathbf{x}}=\frac{1}{K}\sum_{k=1}^{K}\mathbf{x}(k)\mathbf{x}^H(k)$$

§3.3 采样矩阵逆波束形成器

一、SMI 自适应波束形成器

对于由N个阵元组成的传感器阵，假设可以得到K个快拍$\mathbf{x}_1$，$\mathbf{x}_2$，…，$\mathbf{x}_K$，其中$\mathbf{x}_k$为阵列在k时刻得到的N维接收信号矢量。其协方差矩阵为$R_{\mathbf{x}} = E\{\mathbf{x}(k)\mathbf{x}^H(k)\}$。

由§3.2 可知，无结构协方差矩阵的极大似然估计是通过上述快拍得到的采样协方差矩阵：

$$\hat{R}_{\mathbf{x}} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{x}(k)\mathbf{x}^H(k)$$

可见，上述协方差估计实际上是用时间平均代替了协方差中的集平均。

在某些应用中，我们可以得到没有信号时的传感器接收采样，这时可得到噪声协方差矩阵估计：

$$\hat{R}_{\mathbf{n}} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{n}(k)\mathbf{n}^H(k)$$

将上面得到的协方差估计代入到最优波束形成器中，即可得到相应的自适应波束形成器：

$$\hat{\mathbf{w}}^H_{mvdr,smi}=\frac{\mathbf{v}^H\hat{R}_{\mathbf{n}}^{-1}}{\mathbf{v}^H\hat{R}_{\mathbf{n}}^{-1}\mathbf{v}},\qquad \hat{\mathbf{w}}^H_{mpdr,smi}=\frac{\mathbf{v}^H\hat{R}_{\mathbf{x}}^{-1}}{\mathbf{v}^H\hat{R}_{\mathbf{x}}^{-1}\mathbf{v}}$$

$$\hat{\mathbf{w}}^H_{Lcmv,smi}=\hat{R}_{\mathbf{n}}^{-1}\cdot C\left[C^H\hat{R}_{\mathbf{n}}^{-1}C\right]^{-1}\mathbf{f},\qquad \hat{\mathbf{w}}^H_{Lcmp,smi}=\hat{R}_{\mathbf{x}}^{-1}\cdot C\left[C^H\hat{R}_{\mathbf{x}}^{-1}C\right]^{-1}\mathbf{f}$$

$$\hat{\mathbf{w}}^H_{mse,smi}=\hat{R}_{\mathbf{x}}^{-1}\cdot\hat{\mathbf{r}}$$

其中$\hat{\mathbf{r}}=\frac{1}{K}\sum_{k=1}^{K}\mathbf{x}(k)d^*(k)$。

因为上述各式中有$\hat{R}_{\mathbf{n}}^{-1}$或$\hat{R}_{\mathbf{x}}^{-1}$，故被称为采样矩阵逆算法。

二、自适应波束形成算法性能 — 瞬态和稳态性能

一般用阵列输出的信干比来描述统计最优/自适应波束形成器的性能：

$$SINR_0(K)=\frac{\hat{\mathbf{w}}^H(K)\mathbf{v}_s s(K)s^H(K)\mathbf{v}_s^H\hat{\mathbf{w}}(K)}{\hat{\mathbf{w}}^H(K)\left[\mathbf{n}(K)\mathbf{n}^H(K)\right]\hat{\mathbf{w}}(K)}$$

其中K为各传感器接收的快拍的个数，$SINR_0$是一个随机变量，对不同的K，$SINR_0$是随机的，对不同的试验，$SINR_0$亦是随机的。

分析$SINR_0(K)$随K的变化规律，即可知自适应波束形成器的瞬态及稳态性能。

进一步定义一个随机变量$\rho(K)$：

$$\rho(K)=\frac{SINR_0(K)}{SINR_{opt}},\quad 0\le\rho(K)\le 1$$

即可以分析自适应波束形成器相对于相应的最优波束形成器的性能。

对于 **MVDR 波束形成器**，Reed 等人 1974 年给出了$\rho(K)$的统计特性如下：

$$E\{\rho\}=\frac{K+2-N}{K+1},\qquad Var\{\rho\}=\frac{(K+2-N)(N-1)}{(K+1)^2(K+2)}$$

如果期望$E\{SINR_{smi}\}=\alpha\cdot SINR_{mvdr}$，那么需要：

$$K=\frac{1}{1-\alpha}(N-2+\alpha)\approx\frac{1}{1-\alpha}N$$

从上式可见，若希望$\alpha = 0.95$，则需要$K = 20N$。一般在应用中$K = 2N$时的瞬态性能（即$\alpha = 0.5$）就可以达到满意的性能了。

从上式可见，当K从N增加到∞时，$E\{SINR_{mvdr,smi}\}$单调增加，趋于$SINR_{mvdr,opt}$。

对于**MPDR 波束形成器**，94 年 Feldman 和 Griffiths 导出了$E\{SINR_{mpdr,smi}\}$的近似公式为：

$$E\{SINR_{mpdr,smi}\} \approx \frac{SINR_{mpdr} \cdot K}{K + SINR_{mpdr} \cdot (N-1)}$$

可见当K从N增加到∞时，$E\{SINR_{mpdr,smi}\}$亦单调增加趋于$SINR_{mpdr,opt}$。

与自适应 MVDR 波束形成器相比，自适应 MPDR 波束形成器的收敛速度不仅与接收快拍个数有关，还与信噪比有关，信噪比越高，收敛速度越慢。

类似地，如果期望$E\{SINR_{smi}\} = \alpha \cdot SINR_{mvdr}$，那么需要：

$$K = \frac{\alpha}{1-\alpha}(N-1) \cdot SINR_{mpdr} \approx \frac{\alpha}{1-\alpha} N \cdot SINR_{mpdr}$$

若希望$\alpha = 0.95$，则当$SINR_{mpdr} = 10dB$时需$K = 190N$，当$SINR_{mpdr} = 20dB$时$K = 1900N$。若希望$\alpha = 0.5$，则当$SINR_{mpdr} = 10dB$时需要$K = 10N$，当$SINR_{mpdr} = 20dB$时需要$K = 100N$。

对于一般的自适应波束形成器，上述性能的统计特性的解析解、甚至近似解都难以得到。此时，只能用 Monte-Carlo 方法分析自适应波束形成器的瞬态和稳态性能。

§3.4 递推最小二乘波束形成器

递推最小二乘（RLS: Recursive Least Square）波束形成器又可称为递推 SMI 波束形成器。

最小二乘估计方法是高斯在 1795 年研究天体运动时提出的，而最小均方误差估计(MMSE: Minimum Mean Square)是 Kolmogorov 和 Wiener 分别于 1941 和 1949 年独立提出的。在 MMSE 中，所有的二阶统计量均采用集平均来描述，而在 LS 估计中，所有的二阶统计量都用时间平均来表示。

在本小节中，我们以 MPDR 和 MMSE 波束形成器为例，可以看到最小二乘波束形成器与均方意义上的最优波束形成器有同样的形式，只是把计算统计量的集平均换成了时间平均。

一、自适应 MPDR 波束形成器

考虑传感器阵的接收信号：

$$\mathbf{x}(t)=\mathbf{s}+\mathbf{u}=s(t)\mathbf{v}+\sum_{i=1}^{N_u}u_i(t)\mathbf{u}_i+\mathbf{n}$$

波束形成器的输出响应为：

$$y(t)=\mathbf{w}^H\mathbf{x}(t)$$

最优 MPDR 波束形成器的优化目标为：

$$\min_{\mathbf{w}} Var\{y(t)\}=\min_{\mathbf{w}}\mathbf{w}^H R_{\mathbf{x}}\mathbf{w},\quad s.t.\quad \mathbf{w}^H\mathbf{v}=1$$

其中 $R_{\mathbf{x}}=E\{\mathbf{x}(t)\mathbf{x}^H(t)\}$。

在实际应用中，由于上述统计特性 $R_{\mathbf{x}}$ 往往是未知的，可以用如下的最小二乘法来自适应实现 MPDR 波束形成器。

最小二乘算法把上面的目标函数改为使加权输出能量平方和最小：

$$\min_{\mathbf{w}} \varepsilon_y(K) = \min_{\mathbf{w}} \sum_{k=1}^{K} \mu^{K-k} |y(k)|^2, \quad s.t. \quad \mathbf{w}^H(k)\mathbf{v} = 1$$

其中 $y(k) = \mathbf{w}^H(k)\mathbf{x}(k)$，$0 < \mu < 1$是指数加权因子，它允许自适应波束形成器工作在非平稳环境中。

为求解上述最优化问题，令

$$\begin{aligned} J &= \sum_{k=1}^{K} \mu^{K-k} |y(k)|^2 + \lambda\left[\mathbf{w}^H(K)\mathbf{v} - 1 + \mathbf{v}^H\mathbf{w}(K) - 1\right] \\ &= \sum_{k=1}^{K} \mu^{K-k}\mathbf{w}^H(K)\mathbf{x}(k)\mathbf{x}^H(k)\mathbf{w}(K) + \lambda\left[\mathbf{w}^H(K)\mathbf{v} - 1 + \mathbf{v}^H\mathbf{w}(K) - 1\right] \\ &= \mathbf{w}^H(K)\phi(K)\mathbf{w}(K) + \lambda\left[\mathbf{w}^H(K)\mathbf{v} - 1 + \mathbf{v}^H\mathbf{w}(K) - 1\right] \end{aligned}$$

其中 $\phi(k) = \sum_{k=1}^{K} \mu^{K-k}\mathbf{x}(k)\mathbf{x}^H(k)$ 为指数加权采样协方差矩阵。

求 J 对 $\mathbf{w}^H(K)$ 的导数并令之为 0，可得：

$$\hat{\mathbf{w}}_{mpdr}(K)=\Lambda(K)\phi^{-1}(K)\mathbf{v}=\frac{\phi^{-1}(K)\mathbf{v}}{\mathbf{v}^H\phi^{-1}(K)\mathbf{v}}$$

与最优MPDR波束形成器相比，最小二乘MPDR波束形成器用加权输出能量平方和$\varepsilon_y(K)$代替了统计输出能量$Var\{y(t)\}$，用指数加权采样协方差矩阵$\phi(K)$代替了接收信号协方差矩阵$R_{\mathbf{x}}$。

由于$\phi(K)$可以从接收数据$\mathbf{x}(k)$， $k=1,\cdots,K$ 得到，而且$\mu=1$时$\phi(K)$即是$R_{\mathbf{x}}$的极大似然估计$\hat{R}_{\mathbf{x}}$，故最小二乘 MPDR 算法就是自适应 MPDR 波束形成器。$\mu<1$时，相当于在接收数据中加入了遗忘窗，使自适应波束形成器能被用于非平稳环境中。

为了递推实现$\hat{\mathbf{w}}_{mpdr}(K)$，需要从$\phi^{-1}(K-1)$算出$\phi^{-1}(K)$。

$$\phi(K)=\sum_{k=1}^{K}\mu^{K-k}\mathbf{x}(k)\mathbf{x}^H(k)\ =\ \mu\phi(K-1)+\mathbf{x}(K)\mathbf{x}^H(K)$$

由矩阵求逆引理：

$$(A+BCD)^{-1}=A^{-1}-A^{-1}B(C^{-1}+D^{-1}A^{-1}B)^{-1}DA^{-1}$$

可得：

$$\phi^{-1}(K)=\mu^{-1}\phi^{-1}(K-1)-\frac{\mu^{-2}\phi^{-1}(K-1)\mathbf{x}(K)\mathbf{x}^{H}(K)\phi^{-1}(K-1)}{1+\mu^{-1}\mathbf{x}^{H}(K)\phi^{-1}(K-1)\mathbf{x}(K)}$$

令 $P(K)=\phi^{-1}(K)$

及 $\mathbf{g}(K)=\dfrac{\mu^{-1}P(K-1)\mathbf{x}(K)}{1+\mu^{-1}\mathbf{x}^{H}(K)P(K-1)\mathbf{x}(K)}$

则上式变为：

$$P(K)=\mu^{-1}P(K-1)-\mu^{-1}\mathbf{g}(K)\mathbf{x}^{H}(K)P(K-1)$$

$\therefore \quad P(K)\mathbf{x}(K)=\mu^{-1}P(K-1)\mathbf{x}(K)-\mu^{-1}\mathbf{g}(K)\mathbf{x}^{H}(K)P(K-1)\mathbf{x}(K)=\mathbf{g}(K)$

故 $\hat{\mathbf{w}}(K)=\Lambda(K)\phi^{-1}(K)\mathbf{v} \;=\; \Lambda(K)P(K)\mathbf{v} \;=\; \Lambda(K)\left[\mu^{-1}P(K-1)-\mu^{-1}\mathbf{g}(K)\mathbf{x}^{H}(K)P(K-1)\right]\mathbf{v}$

又 $\hat{\mathbf{w}}(K-1)=\Lambda(K-1)P(K-1)\mathbf{v}$

$\therefore \quad \hat{\mathbf{w}}(K)=\mu^{-1}\Lambda(K)\left[I-\mathbf{g}(K)\mathbf{x}^{H}(K)\right]P(K-1)\mathbf{v}$

$$=\mu^{-1}\Lambda(K)\left[I-\mathbf{g}(K)\mathbf{x}^{H}(K)\right]\frac{1}{\Lambda(K-1)}\hat{\mathbf{w}}(K-1)=\frac{\Lambda(K)}{\mu\Lambda(K-1)}\left[I-\mathbf{g}(K)\mathbf{x}^{H}(K)\right]\hat{\mathbf{w}}(K-1)$$

上述 RLS 实现的自适应 MPDR 算法如下：

$P(0)=\frac{1}{\sigma_0^2}I$， $\hat{\mathbf{w}}(0)=\frac{1}{N}\mathbf{v}$，或为任意满足无失真限制的矢量。

对每个快拍 $k=1,\cdots,K$，计算

$$\mathbf{g}(K)=\frac{\mu^{-1}P(K-1)\mathbf{x}(K)}{1+\mu^{-1}\mathbf{x}^H(K)P(K-1)\mathbf{x}(K)}$$

$$P(K)=\mu^{-1}P(K-1)-\mu^{-1}\mathbf{g}(K)\mathbf{x}^H(K)P(K-1)$$

$$\Lambda(K)=\left[\mathbf{v}^H P(K)\mathbf{v}\right]^{-1}$$

$$\hat{\mathbf{w}}_{mpdr}(K)=\frac{\Lambda(K)}{\mu\Lambda(K-1)}\left[I-\mathbf{g}(K)\mathbf{x}^H(K)\right]\hat{\mathbf{w}}_{mpdr}(K-1)=\frac{\Lambda(K)}{\mu\Lambda(K-1)}\left[\hat{\mathbf{w}}_{mpdr}(K-1)-\mathbf{g}(K)\tilde{y}^*(K)\right]$$

其中 $\tilde{y}(K)=\hat{\mathbf{w}}_{mpdr}^H(K-1)\mathbf{x}(K)$。

阵列输出为： $y(K)=\hat{\mathbf{w}}_{mpdr}(K)\mathbf{x}(K)$

二、自适应 MMSE 波束形成器

最优 MMSE 波束形成器的优化目标为：

$$\min_{\mathbf{w}} E\left\{e^2(t)\right\} = \min_{\mathbf{w}} E\left\{\left[d^*(t) - \mathbf{w}^H \mathbf{x}(t)\right]^2\right\}$$

其中 $d(t)$ 为已知的期望响应。

从此得到的波束形成器的权矢量为：

$$\mathbf{w}_{mse} = R_{\mathbf{x}}^{-1}\mathbf{r}$$

其中 $R_{\mathbf{x}} = E\left\{\mathbf{x}(t)\mathbf{x}^H(t)\right\}$，$\mathbf{r} = E\left\{d^*(t)\mathbf{x}(t)\right\}$。

最小二乘波束形成器的优化目标为：

$$\min_{\mathbf{w}} \sum_{k=1}^{K} \mu^{K-k} \left| e(k) \right|^2 = \min_{\mathbf{w}} \sum_{k=1}^{K} \mu^{K-k} \left[d(k) - \mathbf{x}^H(k)\mathbf{w}(K)\right]\left[d^*(k) - \mathbf{w}^H(K)\mathbf{x}(k)\right]$$

求解上述最优化问题，可得到 LS 波束形成器的加权矢量为：

$$\hat{\mathbf{w}}_{ls}(K)=\phi^{-1}(K)\mathbf{r}_{\mathbf{x}d^*}(K)$$

其中 $\phi(K)=\sum_{k=1}^{K}\mu^{K-k}\mathbf{x}(k)\mathbf{x}^H(k)$，$\mathbf{r}_{\mathbf{x}d^*}(K)=\sum_{k=1}^{K}\mu^{K-k}\mathbf{x}(k)d^*(k)$。

上述 LS 波束形成器实际上就是自适应 MMSE 波束形成器。

用与推导递推自适应 MPDR 波束形成器相同的方法，可以得到 RLS 实现的 MMSE 波束形成器。

$$\phi^{-1}(K)=\left[\sum_{k=1}^{K}\mu^{K-k}\mathbf{x}(k)\mathbf{x}^H(k)\right]^{-1}=P(K)\ =\ \mu^{-1}P(K-1)-\mu^{-1}\mathbf{g}(K)\mathbf{x}^H(K)P(K-1)$$

其中 $\mathbf{g}(K)=P(K)\mathbf{x}(K)$。

$$\mathbf{r}_{\mathbf{x}d^*}(K)=\sum_{k=1}^{K}\mu^{K-k}\mathbf{x}(k)d^*(k)\ =\ \mu\cdot\mathbf{r}_{\mathbf{x}d^*}(K-1)+\mathbf{x}(K)d^*(K)$$

$$\therefore\quad \mathbf{w}(K)=\phi^{-1}(K)\mathbf{r}_{\mathbf{x}d^*}(K)\ =P(K)\mathbf{r}_{\mathbf{x}d^*}(K)$$

$$=\left[\mu^{-1}P(K-1)-\mu^{-1}\mathbf{g}(K)\mathbf{x}^H(K)P(K-1)\right]\cdot\left[\mathbf{x}(K)d^*(K)+\mu\cdot\mathbf{r}_{\mathbf{x}d^*}(K-1)\right]$$

$$=\mu^{-1}P(K-1)\mathbf{x}(K)d^*(K)-\mu^{-1}\mathbf{g}(K)\mathbf{x}^H(K)P(K-1)\mathbf{x}(K)d^*(K)$$

$$+P(K-1)\mathbf{r}_{\mathbf{x}d^*}(K-1)-\mathbf{g}(K)\mathbf{x}^H(K)P(K-1)\mathbf{r}_{\mathbf{x}d^*}(K-1)$$

$$=\mathbf{w}(K-1)-\mathbf{g}(K)\mathbf{x}^H(K)\mathbf{w}(K-1)+P(K)\mathbf{x}(K)d^*(K)$$

$$=\mathbf{w}(K-1)+\mathbf{g}(K)\left[d^*(K)-\mathbf{x}^H(K)\mathbf{w}(K-1)\right]$$

故 MMSE 波束形成器的递推自适应算法为：

$P(0)=\dfrac{1}{\sigma_0^2}I$， $\hat{\mathbf{w}}(0)=\dfrac{1}{N}\mathbf{v}$，或为任意满足无失真限制的矢量。

$$\mathbf{g}(K)=\frac{\mu^{-1}P(K-1)\mathbf{x}(K)}{1+\mu^{-1}\mathbf{x}^H(K)P(K-1)\mathbf{x}(K)}$$

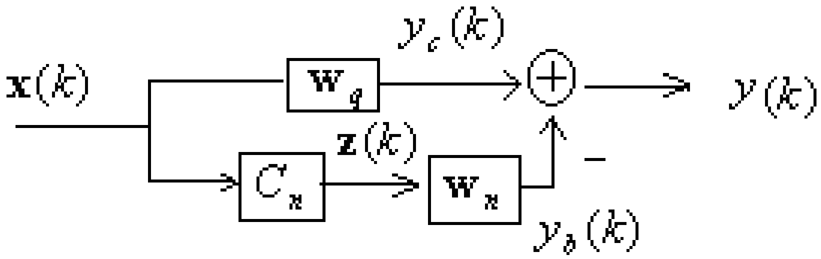
$$P(K)=\mu^{-1}P(K-1)-\mu^{-1}\mathbf{g}(K)\mathbf{x}^H(K)P(K-1)$$

$$\hat{\mathbf{w}}_{mmse}(K)=\hat{\mathbf{w}}_{mmse}(K-1)+\mathbf{g}(K)\left[d^*(K)-\hat{\mathbf{w}}_{mmse}^H(K-1)\mathbf{x}(K)\right]$$

其中 $y(K)=\hat{\mathbf{w}}_{mmse}^H(K)\mathbf{x}(K)$。

三、自适应广义旁瓣对消器

GSC 是 LCMV 的另一种实现方式，它把限制最优化转变为下图所示的无限制最优化问题。



$$\min_{\mathbf{w}_n} Var\{y(t)\} = \min_{\mathbf{w}} E\left\{\left[y_c^*(t) - y_b(t)\right]^2\right\} = \min_{\mathbf{w}} E\left\{\left[y_c^*(t) - \mathbf{w}_n^H \mathbf{z}(t)\right]^2\right\}$$

由此得到的最优 GSC 波束形成器的加权矢量为：

$$\mathbf{w} = \mathbf{w}_q - C_n \mathbf{w}_n = \mathbf{w}_q - C_n (C_n^H R_{\mathbf{x}} C_n)^{-1} C_n^H R_{\mathbf{x}} \mathbf{w}_q$$

由上图可见：$\mathbf{z} = C_n^H \mathbf{x}$，$y_c = \mathbf{w}_q^H \mathbf{x}$

$\therefore \quad R_{\mathbf{z}} = C_n^H R_{\mathbf{x}} C_n$，$R_{\mathbf{z}y_c^*} = C_n^H R_{\mathbf{x}} \mathbf{w}_q$，　其中 $R_{\mathbf{z}} = E\{\mathbf{z} \cdot \mathbf{z}^H\}$。

即最优 GSC 波束形成器的加权矢量为：

$$\mathbf{w} = \mathbf{w}_q - C_n R_{\mathbf{z}}^{-1} R_{\mathbf{z}y_c^*}$$

上式中 $\mathbf{w}_q = C(C^H C)^{-1}\mathbf{f}$ 仅由 LCMV 中的限制确定，与接收数据无关。只有 $\mathbf{w}_n$ 是依赖于数据的，需要自适应。

从前面的分析可见，$\mathbf{w}_n$ 是在使 $y_c(k)$ 与 $y_b(k)$ 之间的均方误差最小意义上得到的权向量。因此，在自适应实现 $\mathbf{w}_n$ 时，可以借用 MMSE 自适应波束形成算法。

$\because$ 最优的 $\mathbf{w}_n$ 为：$\mathbf{w}_n = R_{\mathbf{z}}^{-1} R_{\mathbf{z}y_c^*}$

$\therefore$ 自适应的 $\mathbf{w}_n$ 为：$\hat{\mathbf{w}}_n(K) = \phi_{\mathbf{z}}^{-1}(K)\phi_{\mathbf{z}y_c^*}(K)$

其中 $\phi_{\mathbf{z}}(K) = \sum_{k=1}^{K} \mu^{K-k}\mathbf{z}(k)\mathbf{z}^H(k) = C_n^H \phi_{\mathbf{x}}(K) C_n$

$$\phi_{\mathbf{z}y_c^*}(K)=\sum_{k=1}^{K}\mu^{K-k}\mathbf{z}(k)y_c^*(k)=C_n^H\phi_{\mathbf{x}}(K)\mathbf{w}_q$$

定义 $P_{\mathbf{z}}=\phi_{\mathbf{z}}^{-1}(K)$，则 GSC 波束形成器的 RLS 算法为：

$$P_{\mathbf{z}}(0)=\frac{1}{\sigma_0^2}I$$

$$\hat{\mathbf{w}}_n(0)=\mathbf{0}$$

$$\mathbf{g}_{\mathbf{z}}(K)=\frac{\mu^{-1}P_{\mathbf{z}}(K-1)\mathbf{z}(K)}{1+\mu^{-1}\mathbf{z}^H(K)P_{\mathbf{z}}(K-1)\mathbf{z}(K)}$$

$$P_{\mathbf{z}}(K)=\mu^{-1}P_{\mathbf{z}}(K-1)-\mu^{-1}\mathbf{g}_{\mathbf{z}}(K)\mathbf{z}^H(K)P_{\mathbf{z}}(K-1)$$

$$\hat{\mathbf{w}}_n(K)=\hat{\mathbf{w}}_n(K-1)+\mathbf{g}_{\mathbf{z}}(K)\left[y_c(K)-\hat{\mathbf{w}}_n^H(K-1)\mathbf{z}(K)\right]$$

$$\hat{\mathbf{w}}(K)=\mathbf{w}_q-C_n\hat{\mathbf{w}}_n(K)$$

§3.5 最小均方波束形成器

本小节将介绍两种无限制 MMSE 的最小均方自适应波束形成器和两种线性限制最小均方自适应波束形成器。

第一种无限制 MMSE 最小均方算法是 Stanford 大学的 Widrow 于 1967 年首次提出的 LMS（Least Mean Square）算法，这种算法在过去四十多的自适应阵研究与实现中发挥了重要的作用。

第二种无限制 LMS 算法是 Griffith 于 1969 年提出的，被称为改进的 LMS 算法。

这两种算法均是随机梯度算法，采用梯度技术使均方误差最小。二者的差别仅在于对信号环境假设的先验知识不同。

后两种 LMS 波束形成器分别是针对 LCMV 和 GSC 实现的自适应算法。

一、Widrow 的 LMS 算法

设已知期望的阵列响应$d(t)$，阵列的接收信号为$\mathbf{x}$，阵列波束形成网络中的加权向量为$\mathbf{w}$。

则阵列输出响应与期望响应之间的均方误差为：

$$\xi = E\left\{(d - \mathbf{w}^H \mathbf{x})(d^* - \mathbf{x}^H \mathbf{w})\right\}$$

$$= \sigma_d^2 + \mathbf{w}^H \mathbf{r} - \mathbf{r}^H \mathbf{w} + \mathbf{w} R_{\mathbf{x}} \mathbf{w}$$

其中 $R_{\mathbf{x}} = E\left\{\mathbf{x}\mathbf{x}^H\right\}$，$\mathbf{r} = E\left\{\mathbf{x}d^*\right\}$。

为求解达到上述目标函数的 MMSE 波束形成器的权矢量，对ξ求梯度并令之为 0 得：

$$\nabla \xi_{\mathbf{w}^H} = -\mathbf{r} + R_{\mathbf{x}} \mathbf{w} = 0$$

则最优 MMSE 权向量为：

$$\mathbf{w}_{MSE} = R_{\mathbf{x}}^{-1} \mathbf{r}$$

在本章第二节所介绍的 SMI 算法中，将采样协方差矩阵代入上式即可得自适应算法：

$$\hat{\mathbf{w}}_{MSE} = \hat{R}_{\mathbf{x}}^{-1} \hat{\mathbf{r}}$$

为了避免对$\hat{R}_{\mathbf{x}}$矩阵求逆，第三节利用矩阵求逆引理，介绍了 RLS 算法。

另一种可避免对$\hat{R}_{\mathbf{x}}$矩阵求逆的迭代算法，是采用梯度搜索技术。

当采用上述最优加权向量时，最小均方误差为：

$$\xi_0 = \sigma_d^2 - \mathbf{w}_{MSE}^H R_{\mathbf{x}} \mathbf{w}_{MSE}$$

可见误差表面是二次的，故梯度搜索过程应收敛于唯一的最小点。

定义$\mathbf{w}(K) = \mathbf{w}(K-1) + \alpha(-\nabla \xi_{\mathbf{w}^H})$

其中α是一个实数，被称为步长参数。

将梯度代入上式，得：

$$\mathbf{w}(K) = \mathbf{w}(K-1) + \alpha\left[\mathbf{r} - R_{\mathbf{x}} \mathbf{w}(K-1)\right], \quad K = 1,2,\cdots$$

这就是用确定性梯度算法——最速下降法实现的 MMSE 波束形成器。实际上这种方法并非自适应算法，因为仍要已知$R_{\mathbf{x}}$和$\mathbf{r}$。

在 LMS 算法中，用一种最简单的估计——瞬时值估计出 $R_{\mathbf{x}}$ 和 $\mathbf{r}$ 代替要求已知的二阶统计量 $R_{\mathbf{x}}$ 和 $\mathbf{r}$：

$$\hat{R}_{\mathbf{x}} = \mathbf{x}(K)\mathbf{x}^{H}(K)，\quad \hat{\mathbf{r}}(K) = \mathbf{x}(K)d^{*}(K)$$

这时，梯度的估计值为：

$$\nabla\hat{\xi}_{\mathbf{w}^H} = -\hat{\mathbf{r}} + \hat{R}_{\mathbf{x}}\hat{\mathbf{w}} \quad = \; -\mathbf{x}(K)d^{*}(K) + \mathbf{x}(K)\mathbf{x}^{H}(K)\hat{\mathbf{w}}(K)$$

LMS 算法为：

$$\hat{\mathbf{w}}(K) = \hat{\mathbf{w}}(K-1) + \alpha(-\nabla\hat{\xi}_{\mathbf{w}^H})$$

$$= \; \hat{\mathbf{w}}(K-1) + \alpha\cdot\mathbf{x}(K)[d^{*}(K) - \mathbf{x}^{H}(K)\hat{\mathbf{w}}(K-1)]$$

$$= \; \hat{\mathbf{w}}(K-1) + \alpha\mathbf{x}(K)\left[d^{*}(K) - \hat{y}_{p}^{*}(K)\right] = \; \hat{\mathbf{w}}(K-1) + \alpha\mathbf{x}(K)e_{p}^{*}(K)$$

初值一般可选为 $\hat{\mathbf{w}}(0) = 0$ 或 $\hat{\mathbf{w}}(0) = \dfrac{\mathbf{v}}{N}$。

二、Griffiths 的 LMS 算法

在窄带阵情况，阵列接收数据为：

$$\mathbf{x}(K)=\mathbf{v}_s\cdot s(K)+\mathbf{n}(K)$$

假设已知信号的到达方向（即$\mathbf{v}_s$）以及信号功率σ_s^2。

设$d(K)=s^*(K)$，则：

$$\mathbf{r}=E\{\mathbf{x}(K)d^*(K)\} = E\{\mathbf{x}(K)s^*(K)\}\mathbf{v}_s=\sigma_s^2\cdot\mathbf{v}_s$$

故修正的 LMS 算法为：

$$\hat{\mathbf{w}}(K)=\hat{\mathbf{w}}(K-1)+\alpha(-\nabla\hat{\xi}_{\mathbf{w}^H}(K-1)) = \hat{\mathbf{w}}(K-1)+\alpha(\hat{\mathbf{r}}-\hat{R}_{\mathbf{x}}\hat{\mathbf{w}}(K-1))$$

$$= \hat{\mathbf{w}}(K-1)+\alpha\left[\sigma_s^2\cdot\mathbf{v}_s-\mathbf{x}(K)\mathbf{x}^H(K)\hat{\mathbf{w}}(K-1)\right] = \hat{\mathbf{w}}(K-1)+\alpha\left[\sigma_s^2\cdot\mathbf{v}_s-\mathbf{x}(K)\hat{y}_p^*(K)\right]$$

可见，这种算法不象 Widrow 的 LMS 算法那样需要时间参考信号$d(t)$，但需要知道空间参考信号$\mathbf{v}_s$及信号功率σ_s^2。

三、Frost 的 LMS 算法

对于 LCMP，可以用拉格朗日乘子方法进行线性限制最优化，即：

$$J = \mathbf{w}^H R_{\mathbf{x}} \mathbf{w} + \left[\mathbf{w}^H C - \mathbf{f}^H\right]\lambda + \lambda\left[C^H \mathbf{w} - \mathbf{f}\right]$$

则 $\nabla J_{\mathbf{w}^H} = R_{\mathbf{x}} \cdot \mathbf{w} + C\lambda$

用最速下降法求加权向量，有：

$$\hat{\mathbf{w}}(K) = \hat{\mathbf{w}}(K-1) + \alpha(-\nabla J_{\mathbf{w}^H}) = \hat{\mathbf{w}}(K-1) - \alpha\left[R_{\mathbf{x}}\hat{\mathbf{w}}(K-1) + C\lambda(K-1)\right]$$

同时，$\mathbf{w}(K)$满足下述限制：

$$\mathbf{f} = C^H \mathbf{w}(K) = C^H \mathbf{w}(K-1) - \alpha \cdot C^H R_{\mathbf{x}} \mathbf{w}(K-1) - \alpha \cdot C^H C\lambda(K-1)$$

从中解出入λ，并代入前式，有：

$$\mathbf{w}(K) = \mathbf{w}(K-1) - \alpha\left[I - C(C^H C)^{-1} C^H\right] R_{\mathbf{x}} \mathbf{w}(K-1) + C(C^H C)^{-1}\left[\mathbf{f} - C^H \mathbf{w}(K-1)\right]$$

定义：

$$F=C(C^HC)^{-1}\mathbf{f} \quad \text{及} \quad P_c^{\perp}=I-C(C^HC)^{-1}C^H$$

则 $\mathbf{w}(K)=P_c^{\perp}[\mathbf{w}(K-1)-\alpha R_{\mathbf{x}}\mathbf{w}(K-1)]+F \ =P_c^{\perp}[I-\alpha\cdot R_{\mathbf{x}}]\mathbf{w}(K-1)+F$

$$\mathbf{w}(0)=F=\mathbf{w}_q$$

用瞬时值做为二阶统计量的估计，则 LCMP 波束形成器的 LMS 算法实现为：

$$\hat{\mathbf{w}}(K)=P_c^{\perp}\left[\hat{\mathbf{w}}(K-1)-\alpha\mathbf{x}(K)\mathbf{x}^H(K)\hat{\mathbf{w}}(K-1)\right]+\mathbf{w}_q=\ P_c^{\perp}\left[\hat{\mathbf{w}}(K-1)-\alpha\mathbf{x}(K)\hat{y}_p^*(K)\right]+\mathbf{w}_q$$

对于一般的 LCMP 波束形成器：

$$\mathbf{w}_q=F=C(C^HC)^{-1}\mathbf{f}, \quad P_c^{\perp}=I-C(C^HC)^{-1}C^H$$

而对于 MPDR 波束形成器：

$$\mathbf{w}_q=\frac{1}{N}\mathbf{v}_s, \quad P_c^{\perp}=I-\mathbf{v}_s(\mathbf{v}_s^H\mathbf{v}_s)^{-1}\mathbf{v}_s^H$$

四、广义旁瓣对消器的 LMS 实现

GSC 波束形成器的权矢量为：

$$\mathbf{w} = \mathbf{w}_q - C_n \mathbf{w}_n$$

其中$\mathbf{w}_q$取决于线性限制，而$\mathbf{w}_n$由接收数据决定。

通过求 GSC 上下两个支路的最小均方误差，即可求$\mathbf{w}_n$。

回顾 MMSE-LMS（Widrow）算法：

$$\hat{\mathbf{w}}(K) = \hat{\mathbf{w}}(K-1) + \alpha\mathbf{x}(K)\left[d^*(K) - \hat{y}_p^*(K)\right] = \hat{\mathbf{w}}(K-1) + \alpha\mathbf{x}(K)e_p^*(K)$$

其中 $\hat{y}_p^*(K) = \mathbf{x}^H(K)\hat{\mathbf{w}}(K-1)$

在 GSC 波束形成器中，期望信号为上支路的输出$y_c(t)$，而需要自适应控制的权矢量前的接收信号$\mathbf{z}$是阵列接收信号$\mathbf{x}$经过信号阻塞矩阵得到的，即：

$$\mathbf{z}(K)=C_n\mathbf{x}(K)$$

$$e^*_{PGSC}(K)=y_c(K)-\mathbf{z}^H(K)\hat{\mathbf{w}}^H_n(K-1)$$

$$\therefore\quad \hat{\mathbf{w}}^H_n(K)=\hat{\mathbf{w}}^H_n(K-1)+\alpha\,\mathbf{z}(K)e^*_{PGSC}(K)$$

$$=\hat{\mathbf{w}}^H_n(K-1)+\alpha\,C_n\mathbf{x}(K)e^*_{PGSC}(K)$$

§3.6 几种自适应算法的比较

在上一章的介绍的各种最优波束形成器均可以自适应实现。本章介绍了 MPDR/MVDR、MMSE、LCMP/LCMV（⇔GSC）波束形成器的自适应算法。

在几种自适应算法中，SMI 算法为批处理技术，需求矩阵的逆，其实现自适应的方法为把求二阶统计量的方法由集平均变为时间平均。

RLS、LMS 和 SD 算法均为迭代技术。

SD 算法中假定二阶统计量$R_{\mathbf{x}}$或$R_{\mathbf{n}}$已知，故只是 MMSE 波束形成器的递推算法，而并非真正的自适应算法。

RLS 算法实际上可视为递推 SMI 算法，只是引入了遗忘因子μ。

LMS 算法可以认为是 SD 算法的自适应实现，它用瞬时平均代替了集平均。

$$\hat{\mathbf{w}}_{RLS}(K)=\hat{\mathbf{w}}_{RLS}(K-1)+\phi^{-1}(K)\mathbf{x}(K)\left[d^{*}(K)-\mathbf{x}^{H}(K)\hat{\mathbf{w}}_{RLS}(K-1)\right]$$

$$\hat{\mathbf{w}}_{LMS}(K)=\hat{\mathbf{w}}_{LMS}(K-1)+\alpha\mathbf{x}(K)\left[d^{*}(K)-\mathbf{x}^{H}(K)\hat{\mathbf{w}}_{LMS}(K-1)\right]$$

比较上面两个递推式可见，两种算法的区别仅在于 RLS 中采用$\phi^{-1}(K)=P(K)$，而 LMS 中用α。可见 LMS 是固定步长的迭代算法，而 RLS 是变步长算法。

LMS 算法的特点是计算量小，但其收敛速度受步长α影响，也受$\hat{R}_{\mathbf{x}}$特征值散布程度的影响。

在 RLS 算法中，由于引入了$\phi^{-1}(K)$，相当于把特征值进行了归一化，即去掉了特征值的影响，从而使 RLS 算法的收敛速度与$\hat{R}_{\mathbf{x}}$的特征值散布无关，收敛速度加快。但也同时因为要计算此逆阵而大大增加了计算量。

§3.7 部分自适应波束形成

设传感器的个数为N。则：

RLS 算法每次迭代所需的乘法次数近似为$4N^2$

LMS 算法每次迭代所需的乘法次数近似为$2N$

→ 自适应算法的收敛速度与N有关

部分自适应波束形成只利用部分自由度，只自适应控制一部分的权系数。

降低自由度不仅可以降低计算量，还可以减少自适应算法的瞬态响应时间。

不过，降低自由度一般也会降低性能，一个部分自适应阵往往不能象全自适应阵那样收敛到最优解。

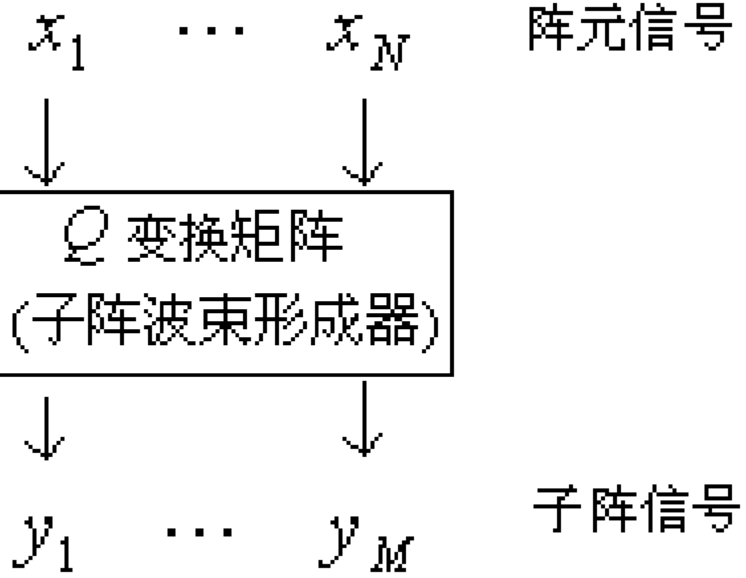
——部分自适应阵的**设计目标**是：在不太降低性能的条件下降低自由度。

一. 自适应子阵波束形成

Chapman 于 76 年论述了自适应子阵波束形成的概念。

如图所示，引入一个$N \times M$维子阵波束形成器变换矩阵Q，把整个阵的N个阵元组元成M个子阵群。

自适应控制每个子阵输出得到总的阵列响应，从而减少信号处理器所需维数。



从图中可得：

$$\mathbf{y} = Q^T \mathbf{x}$$

故子阵接收矢量的协方差矩阵为：

$$R_{\mathbf{y}} = Q^H R_{\mathbf{x}} Q$$

从上一章的分析可知，MVDR 波束形成器的最优加权向量为：

$$\mathbf{w}_{\mathbf{x}} = \beta \cdot R_{\mathbf{x}}^{-1} \cdot \mathbf{v}_{\mathbf{x}}$$

若采用上述子阵接收，则最优加权矢量为：

$$\mathbf{w}_{y_{opt}} = \beta \cdot R_{\mathbf{y}}^{-1} \cdot \mathbf{v}_{\mathbf{y}}$$

其中$\mathbf{v}_{\mathbf{x}}$与$\mathbf{v}_{\mathbf{y}}$ 分别为总阵和子阵的阵簇矢量，二者的关系为：

$$\mathbf{v}_{\mathbf{y}} = Q^T \mathbf{v}_{\mathbf{x}}$$

二、阵元级部分自适应波束形成

自适应子阵波束形成是部分自适应阵问题的一个很好的解决方案，尤其对本来就要求多波束形成的应用更是如此。

然而，实现波束形成矩阵需要额外的开销，如果无需多波束就可以简单地直接控制部分阵元，这就是阵元级部分自适应方案。

这时要解决的问题是：对原阵的哪些阵元进行自适应控制才能使阵性能最优。

为此，需要找出阵元级部分自适应阵的性能指标——如 SINR 或最小总阵输出功率——与阵列构形的关系。在绝大多数情况下，利用上述关系只有计算机才有意义。