阵列信号处理 第七讲

授课教师：段克清
授课单位：电子与通信工程学院
电子邮箱：duankq@mail.sysu.edu.cn
联系电话：xxxxxx（同微信）

再探讨：为什么正则化可以防止过拟合?

正则化：原始损失函数引入额外信息，防止过拟合和提高模型泛化性能的一类方法统称，常对参数取 L_1 范数或 L_2 范数作为额外项。

本质：通过限制参数大小，使模型不能任意拟合训练数据中的随机噪声。

途径：增加惩罚项对复杂模型进行惩罚，强制模型参数值尽可能小，即参数等于 0 或接近于 0，从而约束模型尽量简单。

L_1 正则化的线性回归模型，称为 LASSO 回归。

Least Absolute Shrinkage and Selection Operator。

1996年由 Robert Tibshirani 首次提出，通过构造一个惩罚函数，压缩一些回归系数，从而得到一个较为精炼的模型。

$$\min_{\mathbf{w}}\left[\sum_{i=1}^{N}\left(w_i x_i - y_i\right)^2 + \lambda\left\|\mathbf{w}\right\|_1\right]$$

求解 LASSO 问题的有效方法是最小角度回归

(Least Angle Regressive, LAR 算法)

为什么正则化可以防止过拟合?

L_2 正则化的线性回归模型，称为 Ridge （岭）回归。

Ridge Regression or Tikhonov Regularization

一种改良的最小二乘估计，通过放弃最小二乘法的无偏性，以损失部分信息、降低精度为代价获得回归系数更为符合实际、更可靠的回归方法。对病态数据的拟合要强于最小二乘法。

$$\min_{\mathbf{w}}\left[\sum_{i=1}^{N}\left(w_i x_i - y_i\right)^2 + \lambda\|\mathbf{w}\|_2^2\right]$$

为什么正则化可以防止过拟合?

过拟合特点

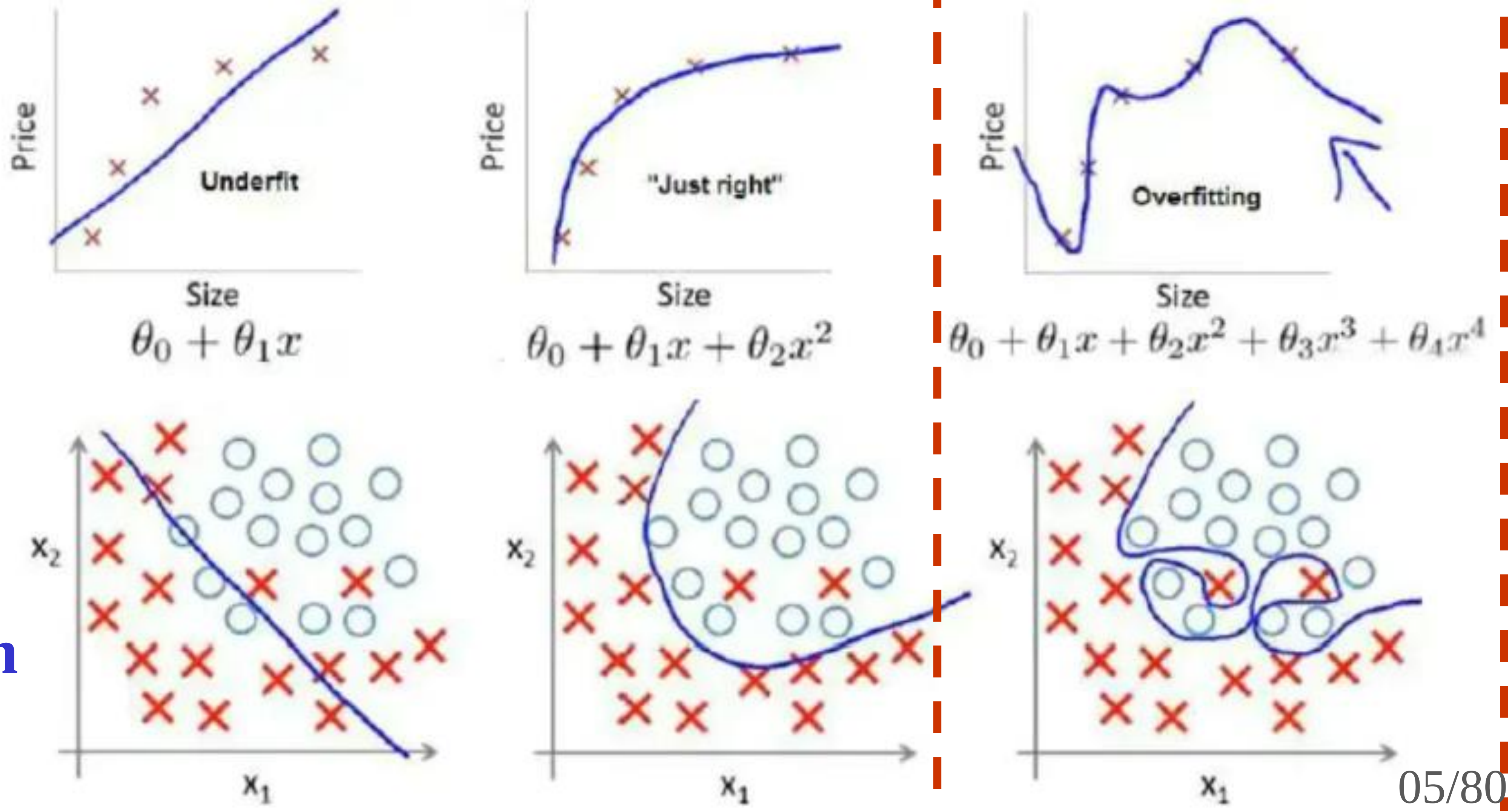
过拟合

回归问题

Regression

分类问题

Classification



为什么正则化可以防止过拟合?

过拟合：学习时选择模型包含参数太多，对当前数据预测很好，对未知数据预测很差。

原因：训练数据有限，如果想在训练数据上表现良好，只能在足够大的模型空间内挑选模型。

奥卡姆剃刀（Ockham’s Razor）原理：如无必要，勿增实体。

应用：在所有可能选择的模型中，能够很好解释已知数据且十分简单才是最好的模型，而正则化可在保证模型有效前提下使得模型简单，且越简单模型泛化能力越强。

如何避免过拟合?

第一个思路：在模型逼近训练数据前提下，尽可能多参数为零（简化模型）。

L_0 范数：模型（参数）稀疏化且易于解释，实现了特征选择；

非连续、非凸、不可导，求解为 NP-Hard 问题。

L_1 范数：模型（参数）稀疏化且易于解释，实现了特征选择；

凸，易求解。

为什么 L_1 范数可以获得稀疏参数？

代价函数： $$\tilde{L}(\mathbf{w}) = L(\mathbf{w}) + \lambda|\mathbf{w}|_1$$

梯度： $$\frac{\partial \tilde{L}(\mathbf{w})}{\partial w_i} = \frac{\partial L(\mathbf{w})}{\partial w_i} + \lambda \operatorname{sgn}(w_i)$$

$$\operatorname{sgn}(w_i) = \begin{cases} 1 & w_i > 0 \\ -1 & w_i < 0 \end{cases}$$

权值更新： $$w_i^{k+1} = w_i^k - \eta\left(\frac{\partial L(\mathbf{w})}{\partial w_i} + \lambda \operatorname{sgn}(w_i)\right)$$

整理后： $$w_i^{k+1} = \left(w_i^k - \lambda \operatorname{sgn}(w_i)\right) - \eta \frac{\partial L(\mathbf{w})}{\partial w_i}$$

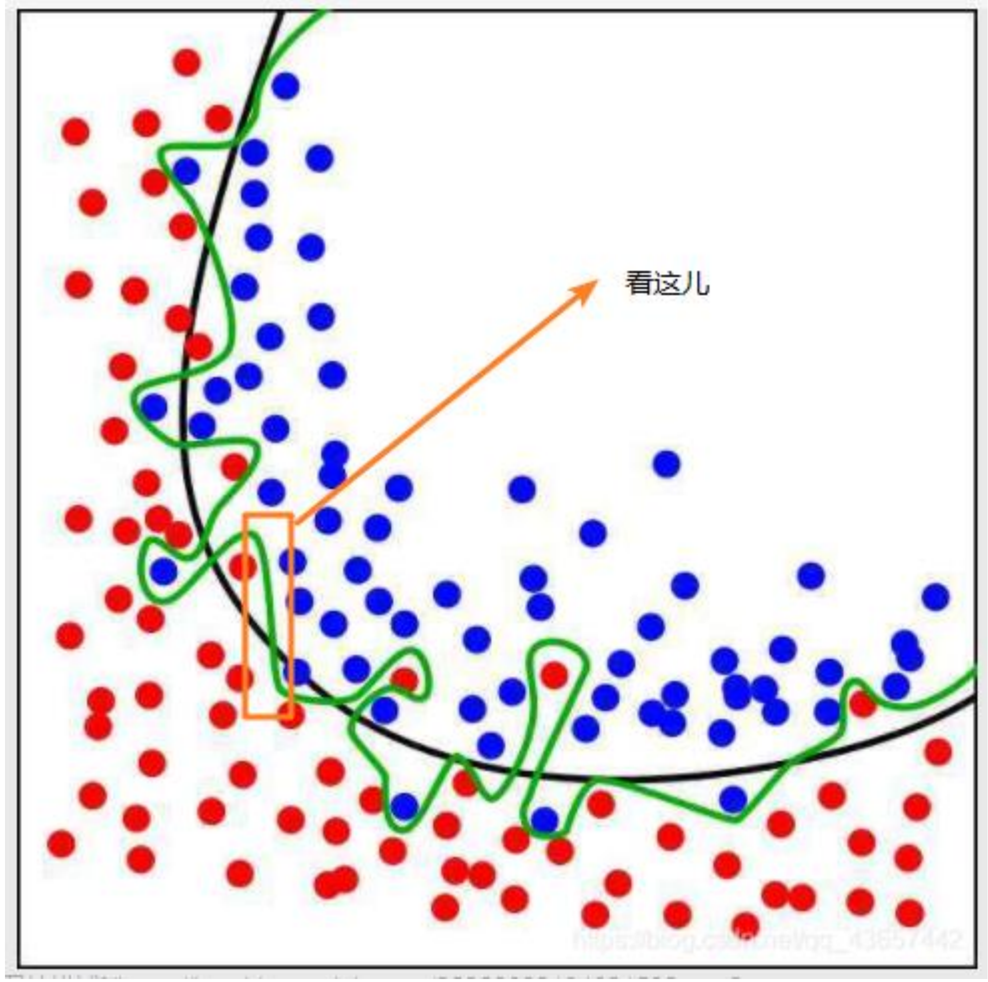
权值每次增减一个常数，更易变为 0

为什么正则化可以防止过拟合？

第二个思路：让所有参数值都尽可能小。

原因：参数值小的模型更能适应不同数据集，在一定程度避免过拟合现象。如线性回归方程，如参数很大，那么只要数据偏移一点点，就会对结果产生很大影响。

L_2 范数：约束权值尽可能小。



变化剧烈，即很多区域偏导很大

为什么 L_2 范数可以获得数值很小的参数（权值）？

代价函数： $$\tilde{L}(\mathbf{w}) = L(\mathbf{w}) + \lambda \left|\mathbf{w}\right|_2^2$$

梯度： $$\frac{\partial \tilde{L}(\mathbf{w})}{\partial w_i} = \frac{\partial L(\mathbf{w})}{\partial w_i} + \lambda w_i$$

权值更新： $$w_i^{k+1} = w_i^k - \eta\left(\frac{\partial L(\mathbf{w})}{\partial w_i} + \lambda w_i\right)$$

整理后： $$w_i^{k+1} = (1-\eta\lambda) w_i^k - \eta \frac{\partial L(\mathbf{w})}{\partial w_i}$$

$\eta\lambda < 1$，因此 w_i 越来越接近 0

L_1 范数和 L_2 范数正则化的区别?

✓ L_1 范数适合于特征选择，适用于只有少数特征起重要作用的场景；L_2 范数只是规则化，适用于大部分特征都起作用的场景。

✓ 正则化的作用就是选择 Loss 与 模型复杂度同时较小的模型，这个最小化像一个下坡的过程，L_1 范数按照绝对值函数的“坡”下降，而 L_2 范数按照二次函数的“坡”下降。

本次课内容

1. L_0范数、L_1范数和L_2范数
2. Tikhonov正则化
3. L_1范数最小化
4. 正交匹配追踪（OMP）法
5. 稀疏贝叶斯学习（SBL）法

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} + \mathbf{n}$$

贝叶斯公式：

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})\,p(\mathbf{x})}{p(\mathbf{y})} \Rightarrow Posterior = \frac{Likelihood \times Prior}{Evidence}$$

Posterior：通过观测数据 y 得到参数 x 的概率，即后验概率；

Likelihood：通过参数 x 得到观测数据 y 的概率；

注: 若 x 确定， y 未知，为概率函数，描述确定参数 x 情况下出现不同数据 y 概率；
若 y 确定，x 未知，为似然函数，描述不同参数 x 情况下出现当前数据 y 的概率。

Prior：参数 x 的先验概率，一般是根据人的先验知识获取。

Evidence：数据 y 发生的概率。 $p(\mathbf{y}) = \int p(\mathbf{y}|\mathbf{x})\,p(\mathbf{x})\,dx$

最大似然估计（MLE） $p\left(\mathbf{y}\middle|\mathbf{x}\right)$ $\mathbf{y}=\mathbf{\Phi}\mathbf{x}+\mathbf{n}$

核心思想： 认为当前发生的事件是概率最大的事件，因此根据给定数据集，寻找使得该数据集发生概率最大的模型参数。

算法特点： 只关注当前获取数据，即只关注当前发生的事情，不考虑事情的先验情况。由于计算简单，且不需要先验知识，因此在机器学习中应用广泛。

最大后验估计（Maximum A Posteriori Probability Estimate, MAP）

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})\,p(\mathbf{x})}{p(\mathbf{y})} \Rightarrow \mathit{Posterior} = \frac{\mathit{Likelihood} \times \mathit{Prior}}{\mathit{Evidence}}$$

核心思想： 允许我们把**先验知识加入到估计模型**中，这在数据很少时特别有用。因为数据很少时我们的观测结果可能出现偏差，此时先验知识可将估计结果“拉”向先验。通过调整先验分布参数，可以调节“拉”向先验的幅度，这些参数称之为“超参数”。

潜在问题： 先验概率选取不当会适得其反。

稀疏贝叶斯学习（Sparse Bayesian Learning, SBL）

SBL 算法是稀疏信号重构的重要方法，由于其不需要设置正则化参数，且性能稳健，因此目前在信号处理领域广泛应用。

该算法涉及高斯分布、最大似然估计、贝叶斯公式、向量求导、期望最大化（EM）算法等数学知识。

下面我们进行详细推导。

$\mathbf{y} = \mathbf{\Phi x} + \mathbf{n}$ ➡ 观测数据 = 字典 × 稀疏系数 + 高斯白噪声

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})\,p(\mathbf{x})}{p(\mathbf{y})}$$ ➡ $$Posterior = \frac{Likelihood \times Prior}{Evidence}$$

目的：通过求取最大后验概率估计求取稀疏系数 x

Step 1：求取似然估计 $p(\mathbf{y}|\mathbf{x})$

Step 2：求取先验概率估计 $p(\mathbf{x})$

Step 3：求取证据 $p(\mathbf{y})$

Step 4：求取后验概率 $p(\mathbf{x}|\mathbf{y})$

Step 5：根据 EM 算法求取 x 中的隐变量（超参数/稀疏系数）

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})\,p(\mathbf{x})}{p(\mathbf{y})}$$

Step 1：求取似然估计 $p(\mathbf{y}|\mathbf{x})$

假定噪声 n 服从均值为 0，方差为 $\sigma^2\mathbf{I}_\mathbf{n}$ 的高斯分布，则可得数据 y 服从

$$p(\mathbf{y}|\mathbf{x}) = \left(2\pi\sigma^2\right)^{-m/2}\exp\left[-\frac{1}{2\sigma^2}\left(\mathbf{y}-\mathbf{\Phi}\mathbf{x}\right)^{\mathrm{H}}\left(\mathbf{y}-\mathbf{\Phi}\mathbf{x}\right)\right]$$

Step 2：求取先验概率 $p(\mathbf{x})$

假定 x 由超参数 γ 产生，并服从均值为 0，方差为 γ_i 的高斯分布，即

$$p(\mathbf{x};\boldsymbol{\gamma}) = (2\pi)^{-\frac{n}{2}}\left|\mathbf{\Gamma}\right|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}\mathbf{x}^{\mathrm{H}}\mathbf{\Gamma}^{-1}\mathbf{x}\right]$$

Step 3：求取证据 $p(\mathbf{y})$

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})\,p(\mathbf{x})}{p(\mathbf{y})}$$

$$p(\mathbf{y};\boldsymbol{\gamma}) = \int_{\mathbf{x}} p(\mathbf{y}|\mathbf{x})\,p(\mathbf{x})\,d\mathbf{x}$$

$$= \int_{\mathbf{x}} \left(2\pi\sigma^2\right)^{-m/2} \left(2\pi\right)^{-\frac{n}{2}} |\boldsymbol{\Gamma}|^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2}(\mathbf{y}-\boldsymbol{\Phi}\mathbf{x})^{\mathrm{H}}(\mathbf{y}-\boldsymbol{\Phi}\mathbf{x}) - \frac{1}{2}\mathbf{x}^{\mathrm{H}}\boldsymbol{\Gamma}^{-1}\mathbf{x}\right] d\mathbf{x}$$

两个高斯函数的卷积仍为高斯函数，将指数函数提出来分析：

$$\mathbf{L} = -\frac{1}{2\sigma^2}(\mathbf{y}-\boldsymbol{\Phi}\mathbf{x})^{\mathrm{H}}(\mathbf{y}-\boldsymbol{\Phi}\mathbf{x}) - \frac{1}{2}\mathbf{x}^{\mathrm{H}}\boldsymbol{\Gamma}^{-1}\mathbf{x}$$

$$= -\frac{1}{2\sigma^2}\left[\mathbf{x}^{\mathrm{H}}\left(\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{\Phi} + \sigma^2\boldsymbol{\Gamma}^{-1}\right)\mathbf{x} - \mathbf{y}^{\mathrm{H}}\boldsymbol{\Phi}\mathbf{x} - \mathbf{x}^{\mathrm{H}}\boldsymbol{\Phi}^{\mathrm{H}}\mathbf{y} + \mathbf{y}^{\mathrm{H}}\mathbf{y}\right]$$

$$\mathbf{L}=-\frac{1}{2\sigma^2}\left[\mathbf{x}^{\mathrm{H}}\left(\mathbf{\Phi}^{\mathrm{H}}\mathbf{\Phi}+\sigma^2\mathbf{\Gamma}^{-1}\right)\mathbf{x}-\mathbf{y}^{\mathrm{H}}\mathbf{\Phi}\mathbf{x}-\mathbf{x}^{\mathrm{H}}\mathbf{\Phi}^{\mathrm{H}}\mathbf{y}+\mathbf{y}^{\mathrm{H}}\mathbf{y}\right]$$

L 是关于 x 的二次项。对于高斯函数，具有以下性质：

$$\int_{\mathbf{a}}\exp\left[-\left(\mathbf{b}-\mathbf{\Phi}\mathbf{a}\right)^{\mathrm{H}}\left(\mathbf{b}-\mathbf{\Phi}\mathbf{a}\right)\right]d\mathbf{a}=C$$

为方便求积分，我们需将 L 表达为 $-\left(\mathbf{y}-\mathbf{\Phi}\mathbf{x}\right)^{\mathrm{H}}\left(\mathbf{y}-\mathbf{\Phi}\mathbf{x}\right)+f\left(\mathbf{y},\sigma^2\right)$ **形式**

思路： **前项** $-\left(\mathbf{y}-\mathbf{\Phi}\mathbf{x}\right)^{\mathrm{H}}\left(\mathbf{y}-\mathbf{\Phi}\mathbf{x}\right)=0$ **等于 0 时，可求得** $f\left(\mathbf{y},\sigma^2\right)$

求 x，使得 $-\left(\mathbf{y}-\mathbf{\Phi}\mathbf{x}\right)^{\mathrm{H}}\left(\mathbf{y}-\mathbf{\Phi}\mathbf{x}\right)=0$

然后将该 x 代入 L 中，就可以得到 $f\left(\mathbf{y},\sigma^2\right)$ **的形式。**

如何求 x？ $$\mathbf{L}=-\frac{1}{2\sigma^2}\left[\mathbf{x}^{\mathrm{H}}\left(\mathbf{\Phi}^{\mathrm{H}}\mathbf{\Phi}+\sigma^2\mathbf{\Gamma}^{-1}\right)\mathbf{x}-\mathbf{y}^{\mathrm{H}}\mathbf{\Phi}\mathbf{x}-\mathbf{x}^{\mathrm{H}}\mathbf{\Phi}^{\mathrm{H}}\mathbf{y}+\mathbf{y}^{\mathrm{H}}\mathbf{y}\right]$$

满足 $$-\left(\mathbf{y}-\mathbf{\Phi}\mathbf{x}\right)^{\mathrm{H}}\left(\mathbf{y}-\mathbf{\Phi}\mathbf{x}\right)=0$$

那么 $$\mathbf{L}=-\left(\mathbf{y}-\mathbf{\Phi}\mathbf{x}\right)^{\mathrm{H}}\left(\mathbf{y}-\mathbf{\Phi}\mathbf{x}\right)+f\left(\mathbf{y},\sigma^2\right)$$ **与 x 无关**

那么 $$\frac{d\mathbf{L}}{d\mathbf{x}}=\frac{d\left\{-\frac{1}{2\sigma^2}\left[\mathbf{x}^{\mathrm{H}}\left(\mathbf{\Phi}^{\mathrm{H}}\mathbf{\Phi}+\sigma^2\mathbf{\Gamma}^{-1}\right)\mathbf{x}-\mathbf{y}^{\mathrm{H}}\mathbf{\Phi}\mathbf{x}-\mathbf{x}^{\mathrm{H}}\mathbf{\Phi}^{\mathrm{H}}\mathbf{y}+\mathbf{y}^{\mathrm{H}}\mathbf{y}\right]\right\}}{d\mathbf{x}}=0$$

得到 $$\mathbf{x}=\left(\mathbf{\Phi}^{\mathrm{H}}\mathbf{\Phi}+\sigma^2\mathbf{\Gamma}^{-1}\right)^{-1}\mathbf{\Phi}^{\mathrm{H}}\mathbf{y}$$

得到 $$\mathbf{x}=\left(\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{\Phi}+\sigma^{2}\boldsymbol{\Gamma}^{-1}\right)^{-1}\boldsymbol{\Phi}^{\mathrm{H}}\mathbf{y}$$

将 x 代入到 L 中可得

$$f\left(\mathbf{y},\sigma^{2}\right)=\mathbf{L}=-\frac{1}{2\sigma^{2}}\mathbf{y}^{\mathrm{H}}\left[\mathbf{I}-\boldsymbol{\Phi}\left(\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{\Phi}+\sigma^{2}\boldsymbol{\Gamma}^{-1}\right)^{-1}\boldsymbol{\Phi}^{\mathrm{H}}\right]\mathbf{y}$$

对证据进行积分

$$\begin{aligned}p(\mathbf{y};\boldsymbol{\gamma})&=C\exp\left[f\left(\mathbf{y},\sigma^{2}\right)\right]\\&=C\exp\left\{-\frac{1}{2\sigma^{2}}\mathbf{y}^{\mathrm{H}}\left[\mathbf{I}-\boldsymbol{\Phi}\left(\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{\Phi}+\sigma^{2}\boldsymbol{\Gamma}^{-1}\right)^{-1}\boldsymbol{\Phi}^{\mathrm{H}}\right]\mathbf{y}\right\}\end{aligned}$$

$$p(\mathbf{y};\boldsymbol{\gamma}) = C\exp\left\{-\frac{1}{2\sigma^2}\mathbf{y}^{\mathrm{H}}\left[\mathbf{I}-\mathbf{\Phi}\left(\mathbf{\Phi}^{\mathrm{H}}\mathbf{\Phi}+\sigma^2\mathbf{\Gamma}^{-1}\right)^{-1}\mathbf{\Phi}^{\mathrm{H}}\right]\mathbf{y}\right\}$$

可看出 $p(\mathrm{y};\gamma)$ 是服从高斯分布，其均值为 0，协方差矩阵 $\sum_{\mathbf{y}}^{-1}$ 满足

$$\sum\nolimits_{\mathbf{y}}^{-1} = \frac{1}{\sigma^2}\left[\mathbf{I}-\mathbf{\Phi}\left(\mathbf{\Phi}^{\mathrm{H}}\mathbf{\Phi}+\sigma^2\mathbf{\Gamma}^{-1}\right)^{-1}\mathbf{\Phi}^{\mathrm{H}}\right]$$

根据如下矩阵求逆公式

$$\left(\mathbf{A}+\mathbf{BCD}\right)^{-1} = \mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{BC}\left(\mathbf{I}+\mathbf{DA}^{-1}\mathbf{BC}\right)^{-1}\mathbf{DA}^{-1}$$

求得

$$\sum\nolimits_{\mathbf{y}} = \sigma^2\mathbf{I}+\mathbf{\Phi\Gamma\Phi}^{\mathrm{H}}$$

Step 4 :求取后验概率 $p(\mathbf{x}|\mathbf{y})$

$$p(\mathbf{x}|\mathbf{y};\boldsymbol{\gamma})=\frac{p(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{y};\boldsymbol{\gamma})}$$

$$=C\frac{\exp\left[-\frac{1}{2\sigma^2}(\mathbf{y}-\boldsymbol{\Phi}\mathbf{x})^{\mathrm{H}}(\mathbf{y}-\boldsymbol{\Phi}\mathbf{x})\right]\bullet\exp\left[-\frac{1}{2}\mathbf{x}^{\mathrm{H}}\boldsymbol{\Gamma}^{-1}\mathbf{x}\right]}{\exp\left\{-\frac{1}{2\sigma^2}\mathbf{y}^{\mathrm{H}}\left[\mathbf{I}-\boldsymbol{\Phi}\left(\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{\Phi}+\sigma^2\boldsymbol{\Gamma}^{-1}\right)^{-1}\boldsymbol{\Phi}^{\mathrm{H}}\right]\mathbf{y}\right\}}$$

$$=C\exp\left[-\frac{1}{2\sigma^2}\left[-\mathbf{y}^{\mathrm{H}}\boldsymbol{\Phi}\mathbf{x}-\mathbf{x}^{\mathrm{H}}\boldsymbol{\Phi}^{\mathrm{H}}\mathbf{y}+\mathbf{x}^{\mathrm{H}}\mathbf{x}+\mathbf{y}^{\mathrm{H}}\left(\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{\Phi}+\sigma^2\boldsymbol{\Gamma}^{-1}\right)^{-1}\mathbf{y}\right]-\frac{1}{2}\mathbf{x}^{\mathrm{H}}\boldsymbol{\Gamma}^{-1}\mathbf{x}\right]$$

目的： 求取 x 的协方差矩阵 Γ（其对角线元素为稀疏系数），因此无法直接利用最大后验概率求解。

特点： y 为观测数据集，x 为隐变量集，估计参数 Γ 。

期望最大化（Expectation Maximization ）法

- ✓ 初始化一个 Γ^0 参数;
- ✓ **E-Step：** 根据 Γ^k 计算出 x 对数似然函数，然后求期望值（此处用到最大后验概率）；
- ✓ **M-Step：** 求该期望最大化得到参数 Γ^{k+1}；
- ✓ 重复上述 EM 步骤直至收敛。

E-Step：

最大后验概率估计为：两个高斯概率密度的乘积再除以一个高斯概率密度。因此，其结果仍为高斯分布。

$$f(x)=\frac{1}{\sqrt{2\pi\sigma}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

均值为指数部分对 x 的一阶导数取零 ⇒ $\boldsymbol{\mu}_{\mathbf{x}}=\left(\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{\Phi}+\sigma^2\boldsymbol{\Gamma}^{-1}\right)^{-1}\boldsymbol{\Phi}^{\mathrm{H}}\mathbf{y}$

协方差矩阵的逆为指数部分对 x 的二阶导数 ⇒ $\sum_{\mathbf{x}}^{-1}=\frac{1}{\sigma^2}\boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{H}}+\boldsymbol{\Gamma}^{-1}$

由于 $m \ll n$ ，所以 $\Sigma_{\mathbf{x}}^{-1}$ 求逆的复杂度远高于 $\Sigma_{\mathbf{y}}^{-1}$ 求逆复杂度，所以运用矩阵和求逆公式进一步得到：

$$\boldsymbol{\mu}_{\mathbf{x}} = \boldsymbol{\Gamma}\boldsymbol{\Phi}^{\mathrm{H}}\Sigma_{\mathbf{y}}^{-1}\mathbf{y}$$

$$\Sigma_{\mathbf{x}} = \boldsymbol{\Gamma} - \boldsymbol{\Phi}^{\mathrm{H}}\Sigma_{\mathbf{y}}^{-1}\boldsymbol{\Phi}\boldsymbol{\Gamma}$$

M-Step：

$$\boldsymbol{\gamma}_n^{new} = \|\boldsymbol{\mu}_{\mathbf{x}}\|_2^2 + \sum\nolimits_{\mathbf{x}}(n,n)$$

$$\left(\sigma^2\right)^{new} = \frac{\|\mathbf{y}-\boldsymbol{\Phi}\boldsymbol{\mu}_{\mathbf{x}}\|_2^2 + \left(\sigma^2\right)^{old}\sum_{n=1}^{N}\left[1-\sum\nolimits_{\mathbf{x}}(n,n)\big/\boldsymbol{\gamma}_n^{old}\right]}{M}$$

更新 $\sum\nolimits_{\mathbf{y}} = \sigma^2\mathbf{I} + \boldsymbol{\Phi}\boldsymbol{\Gamma}\boldsymbol{\Phi}^{\mathrm{H}}$ **代入 E-Step 多次迭代，直到收敛。**

https://www.jianshu.com/p/1168682b87f2?utm_campaign=haruki

阵列信号处理 第八讲

授课教师：段克清

授课单位：电子与通信工程学院

电子邮箱：duankq@mail.sysu.edu.cn

联系电话：xxxxxxx（同微信）

本次课内容

1. ADBF基本概念
2. 常规波束形成
3. 自适应波束形成

本次课内容

1. ADBF基本概念
2. 常规波束形成
3. 自适应波束形成

自适应阵列处理、空域自适应滤波、自适应数字波束形成

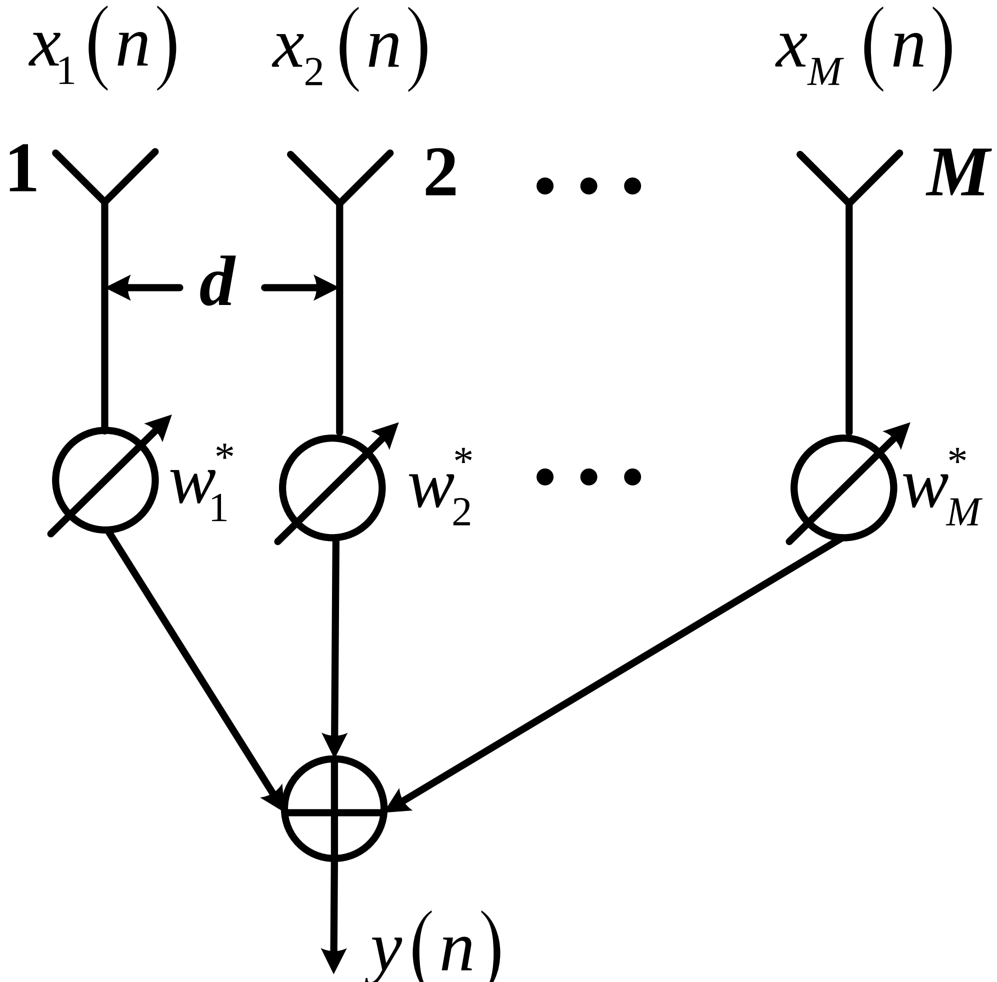
- 自适应阵列处理，也称空域自适应滤波。其通过一定布置的空间阵元对空间信号场进行采样，然后经加权相加处理得到期望输出结果。
- 自适应阵列处理对特定信号的接收和对干扰的抑制均通过形成自适应方向图来实现，因此自适应阵列处理也称为自适应数字波束形成（Adaptive Digital BeamForming，ADBF）

ADBF的核心目的

通过调整各阵元权值，实现对有用信号的有效接收。

什么是有效接收?

- 使阵列方向图主瓣（阵列天线最大增益方向）对准期望信号；
- 对干扰进行有效抑制。

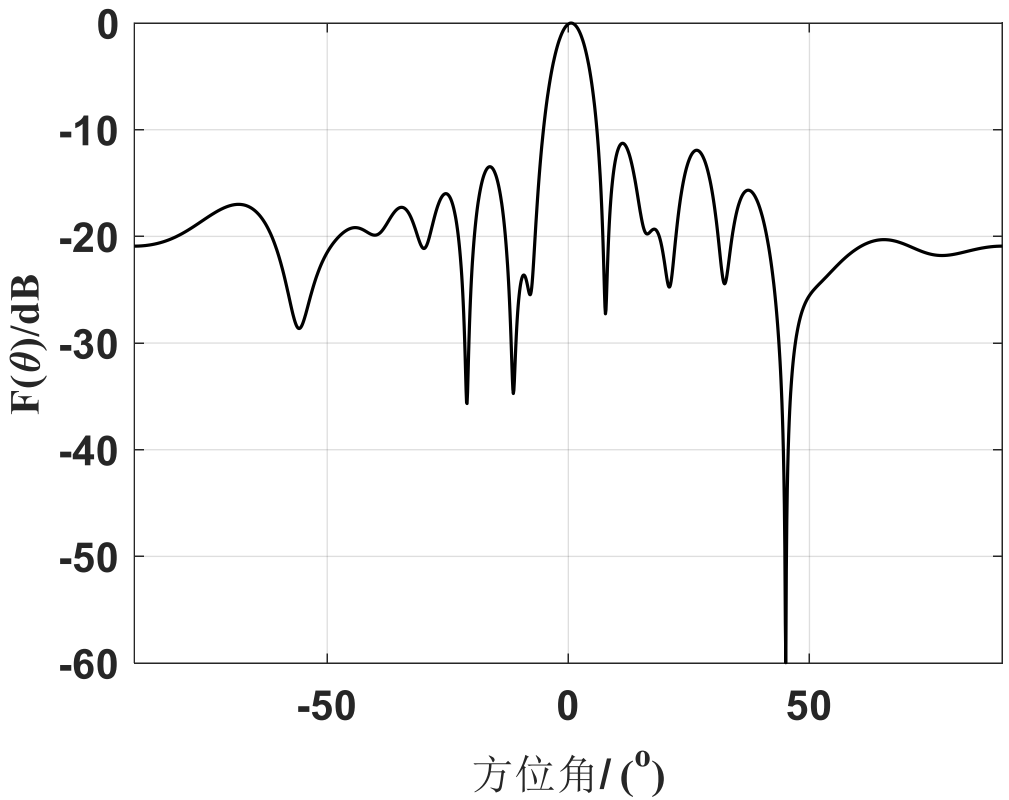


波束形成：用一定形状的波束来通过有用信号或需要方向的信号，并抑制不需要方向的干扰；

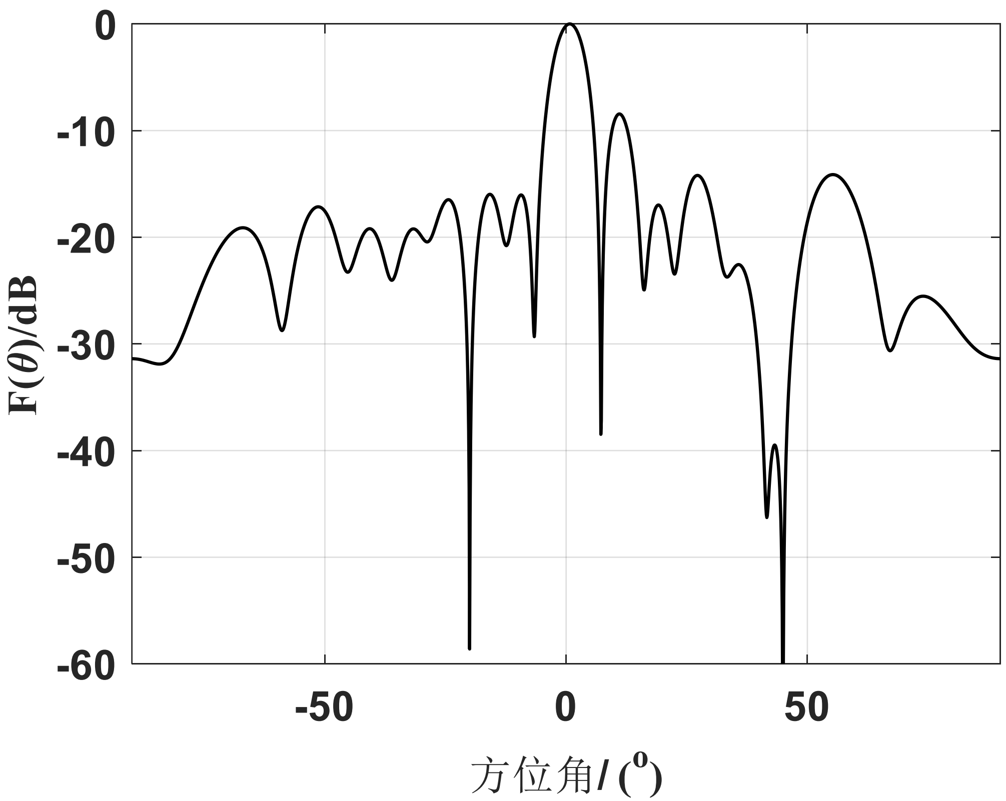
阵列天线波束形成可采用模拟方式，也可采用数字方式。采用数字方式在基带实现滤波的技术称为数字波束形成（DBF），是空域滤波主要形式，在通信中也称之为智能天线。

自适应方向图

主波束方向（目标方向）： 0°



干扰方向：45°



干扰方向：[-20° 45°]

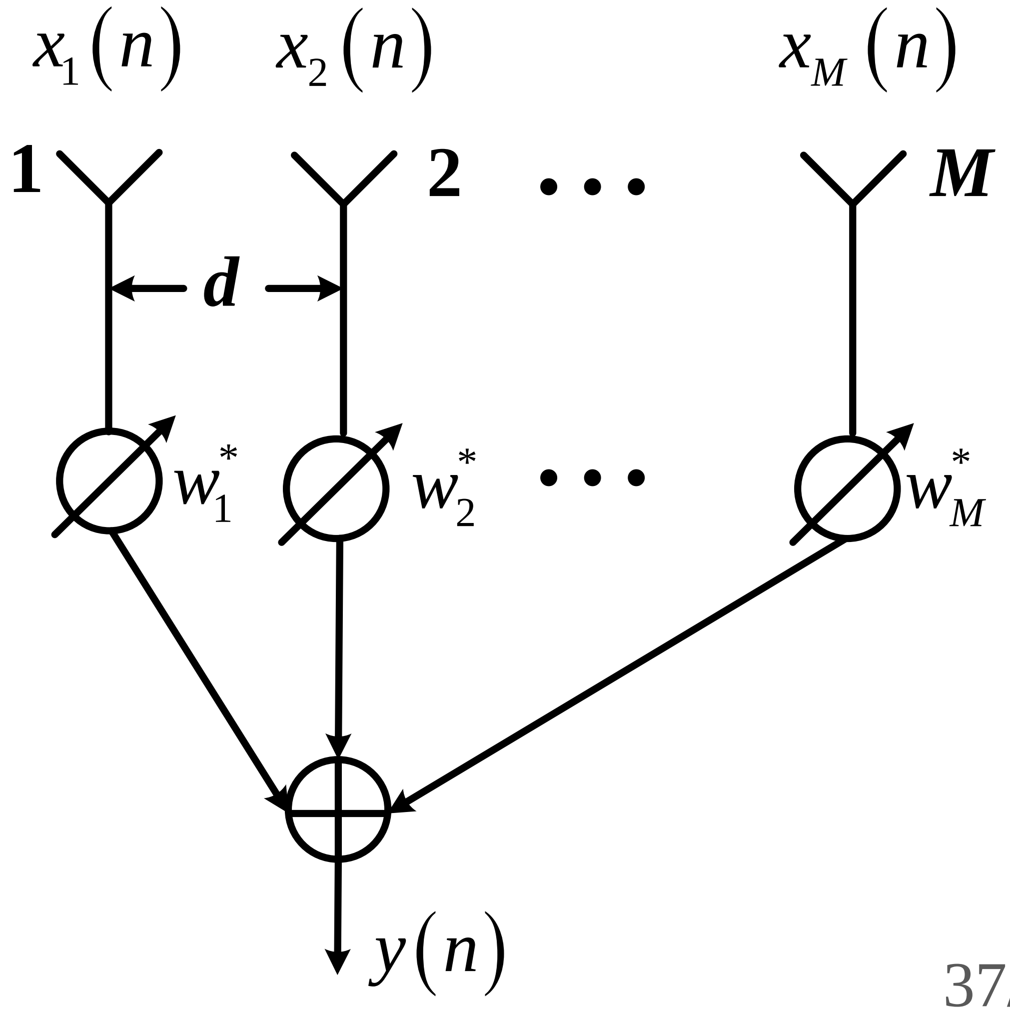
本次课内容

1. ADBF基本概念
2. 常规波束形成
3. 自适应波束形成

常规波束形成（Conventional Beamforming, CBF）

窄带条件：

同时刻采集数据，所有阵元上信号的复包络相同，只需考虑相位变化，而它只依赖于阵列的几何结构。对于等距线阵，则更简单，只依赖于与 x 轴的夹角。



不考虑噪声，阵列接收信号模型为

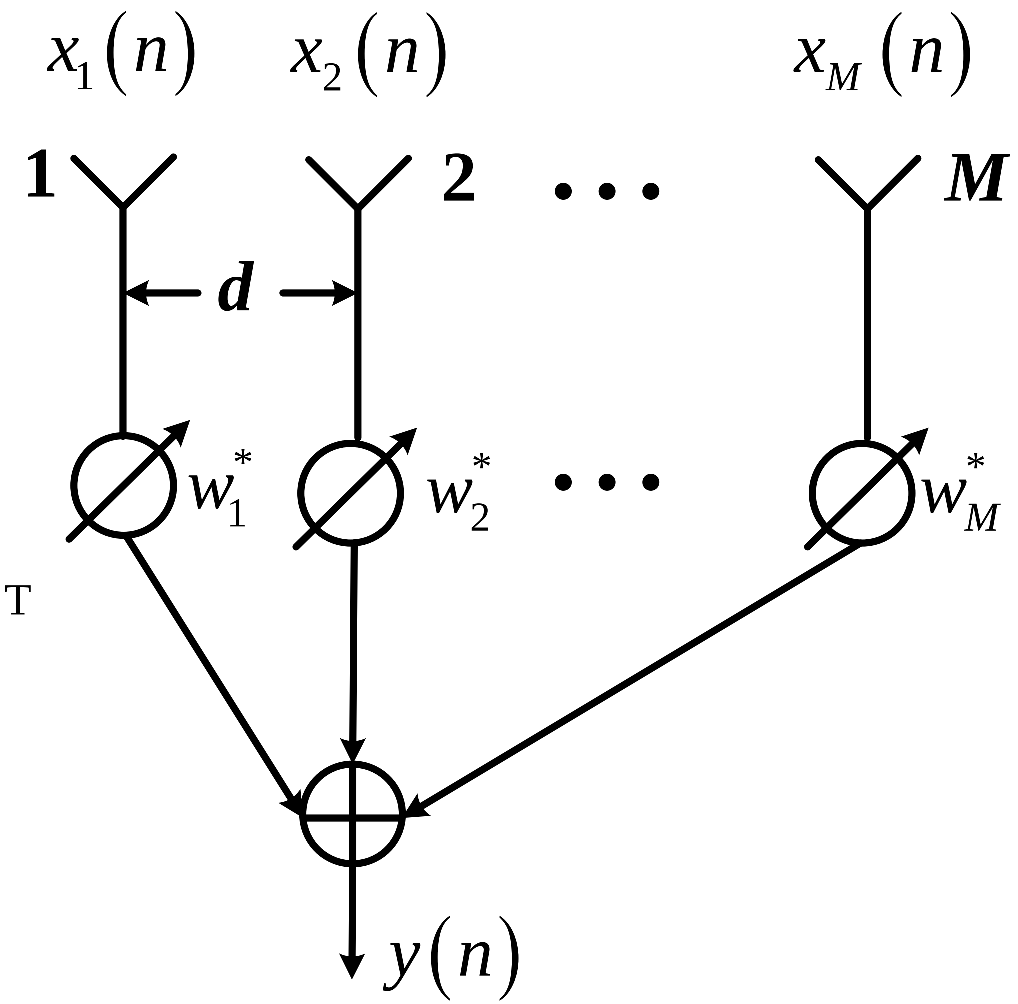
$$\mathbf{x}(n)=\mathbf{a}(\theta)\mathbf{s}(n)\quad n=1,2,\cdots,N$$

其中 $\mathbf{x}(n)\in\mathbb{C}^{M\times 1};\ \mathbf{a}(\theta)\in\mathbb{C}^{M\times 1};$

$\mathbf{s}(n)\in\mathbb{C}^{P\times 1};$

$$\mathbf{a}(\theta)=\begin{bmatrix}1 & e^{-j\frac{2\pi d\sin\theta}{\lambda}} & \cdots & e^{-j\frac{2\pi(M-1)d\sin\theta}{\lambda}}\end{bmatrix}^{\mathrm{T}}$$

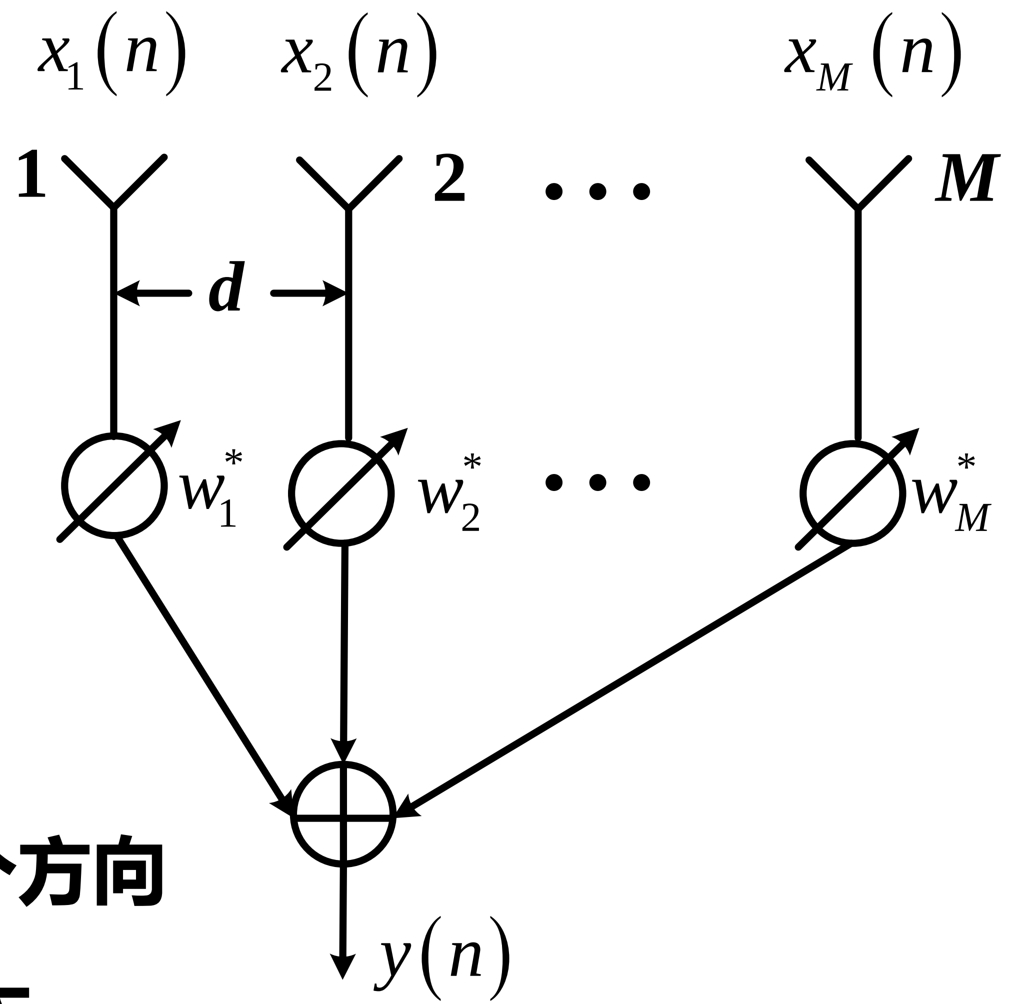
为导向矢量（Steering Vector）



Beamforming 基本思想

通过将各阵元输出进行加权求和，在同一时间将天线波束“导向”到期望方向上。阵列输出可表示为

$$\mathbf{y}(n)=\mathbf{w}^{\mathrm{H}}(n)\mathbf{x}(n)=\mathbf{w}^{\mathrm{H}}\mathbf{a}(\theta)\mathbf{s}(n)$$



目的：增强特定方向信号的功率。

记 $F(\theta)=\mathbf{w}^{\mathrm{H}}\mathbf{a}(\theta)$ 为方向图。当 w 对某个方向 θ_0 的信号同相相加时得到 $F(\theta)$ 的模值最大。

阵列的方向图

阵列输出的绝对值与来波方向之间的关系称为天线的方向图。

方向图一般分两类：

- 静态方向图：阵列输出直接相加（不考虑信号及来波方向），其阵列的最大值出现在阵列法线方向（即 $\theta = 0^o$）
- 带指向的方向图：信号的指向通过控制加权相位来实现，即常说的相控阵列。

对于 x(n) 实际上是空域采样信号，波束形成实现了对 θ 的选择，即实现空域滤波。这一点可对比时域滤波，实现频率选择。

等距线阵情况：

若要波束形成指向 θ_0，则可取 $\mathbf{w} = \mathbf{a}(\theta_0)$ 。波束形成：

$$Y_0 = \mathbf{a}^{\mathrm{H}}(\theta_0)\mathbf{a}(\theta)$$

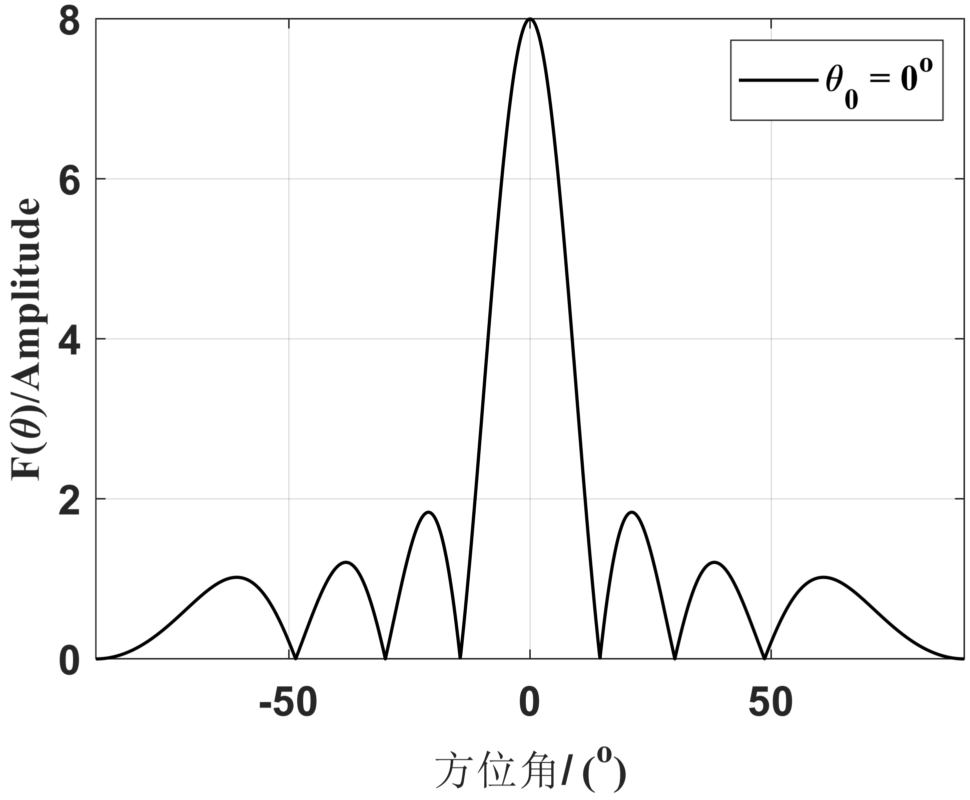
$$= Mg_0 e^{-j\frac{2\pi(M-1)d(\sin\theta_0 - \sin\theta)}{\lambda}} \frac{\sin\left[\pi Md\left(\sin\theta_0 - \sin\theta\right)/\lambda\right]}{M\sin\left[\pi d\left(\sin\theta_0 - \sin\theta\right)/\lambda\right]}$$

方向图为： $$F(\theta)=\left|\frac{\sin\left[\pi Md\left(\sin\theta_0-\sin\theta\right)/\lambda\right]}{M\sin\left[\pi d\left(\sin\theta_0-\sin\theta\right)/\lambda\right]}\right|$$

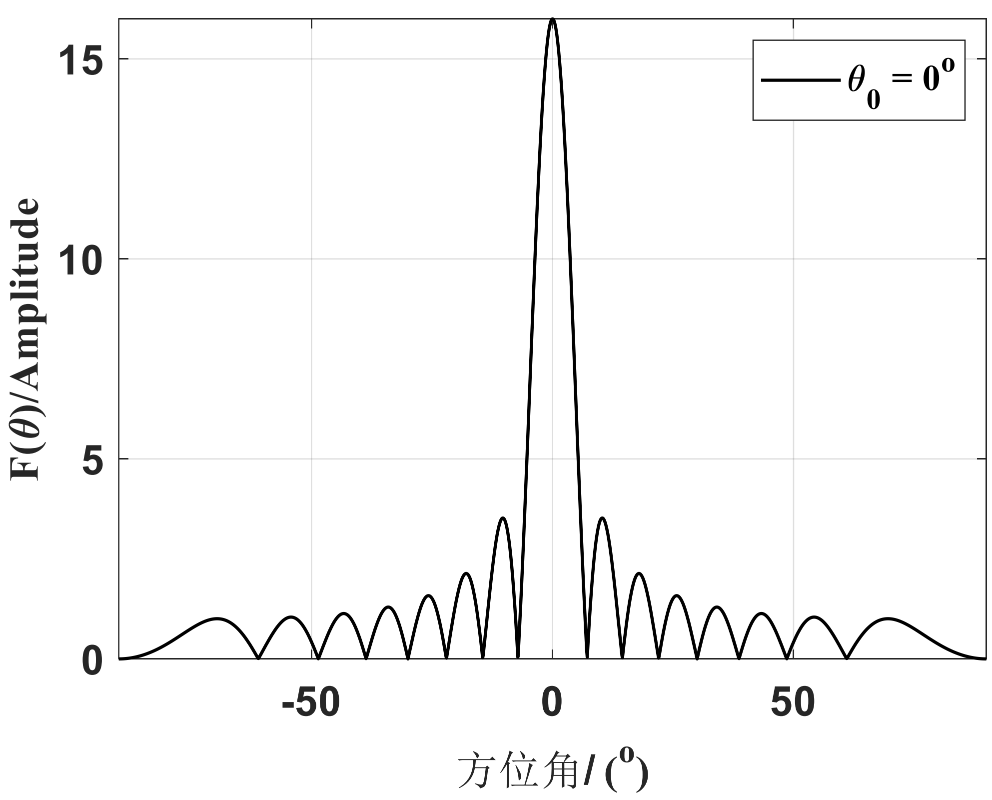
该方向图由如下特点：

- **波束成 $sinx/x$形状，其最大值为 M（注意方向图未归一化）。波瓣半功率点宽度为 $\theta_{0.5}=0.886\lambda/Md\,(rad)=50.8\lambda/Md\,$(度)。根据主瓣宽度正比于天线孔径的倒数。**
- **最大副瓣为第一副瓣，且约-13.4dB。这种副瓣电平对于很多应用来说太大了，为降低副瓣必须采用幅度加权（又称为加窗）。**

幅度方向图（未归一化）

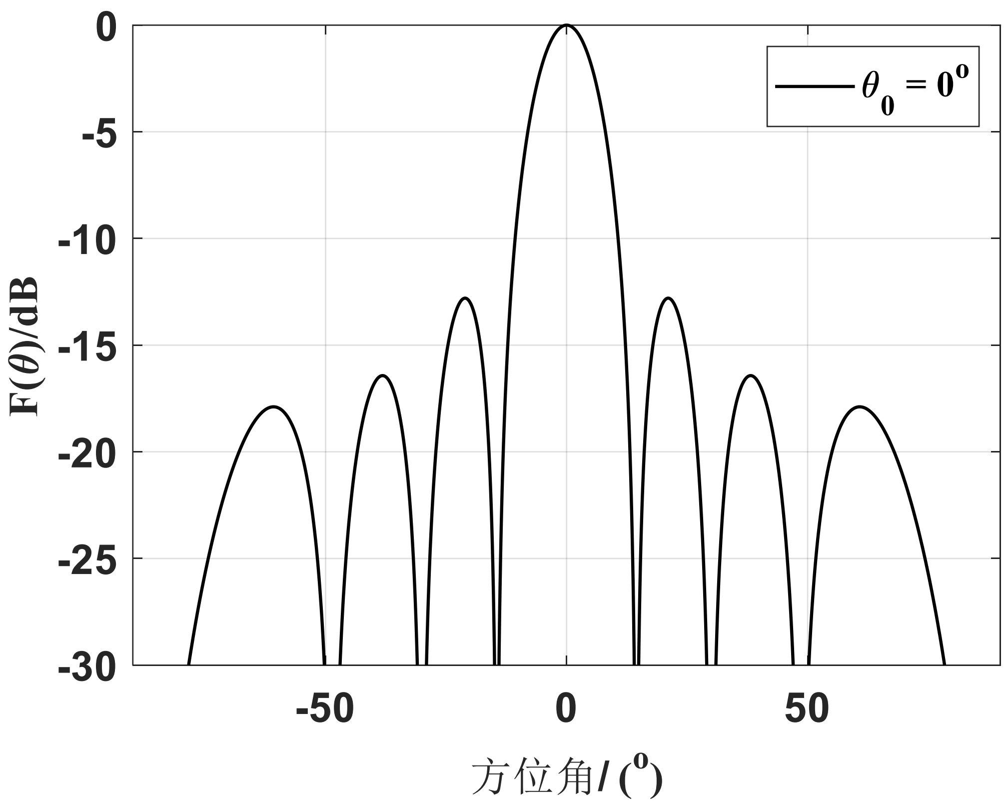


$M = 8$

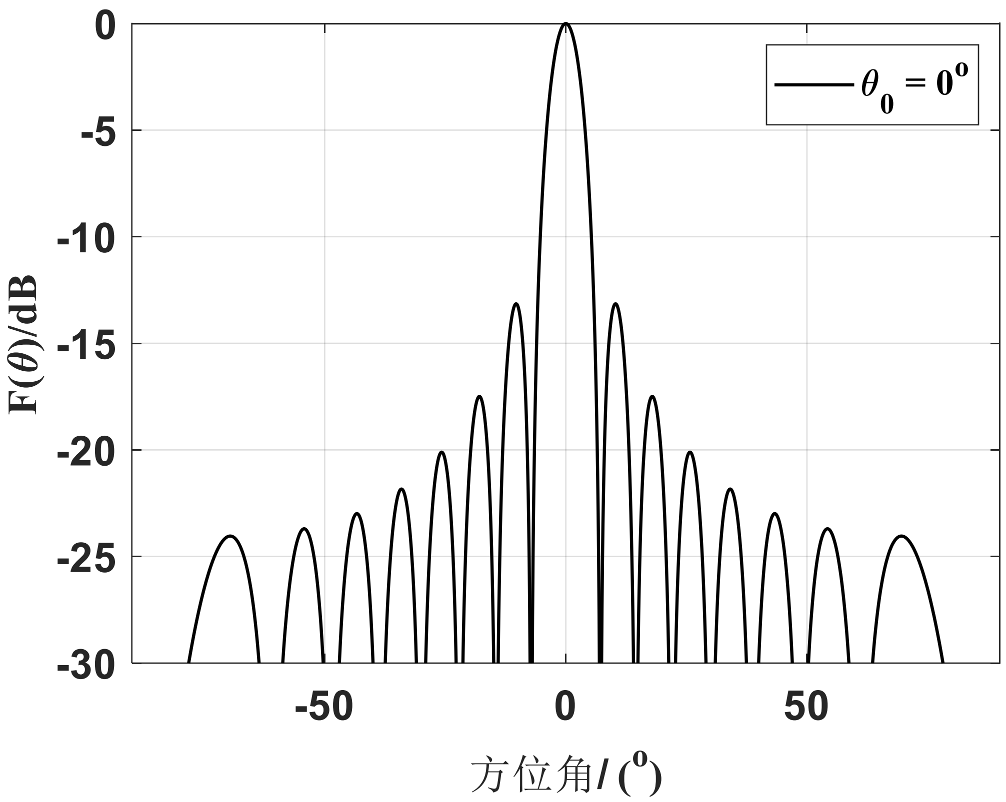


$M = 16$

归一化功率方向图



$M = 8$



$M = 16$

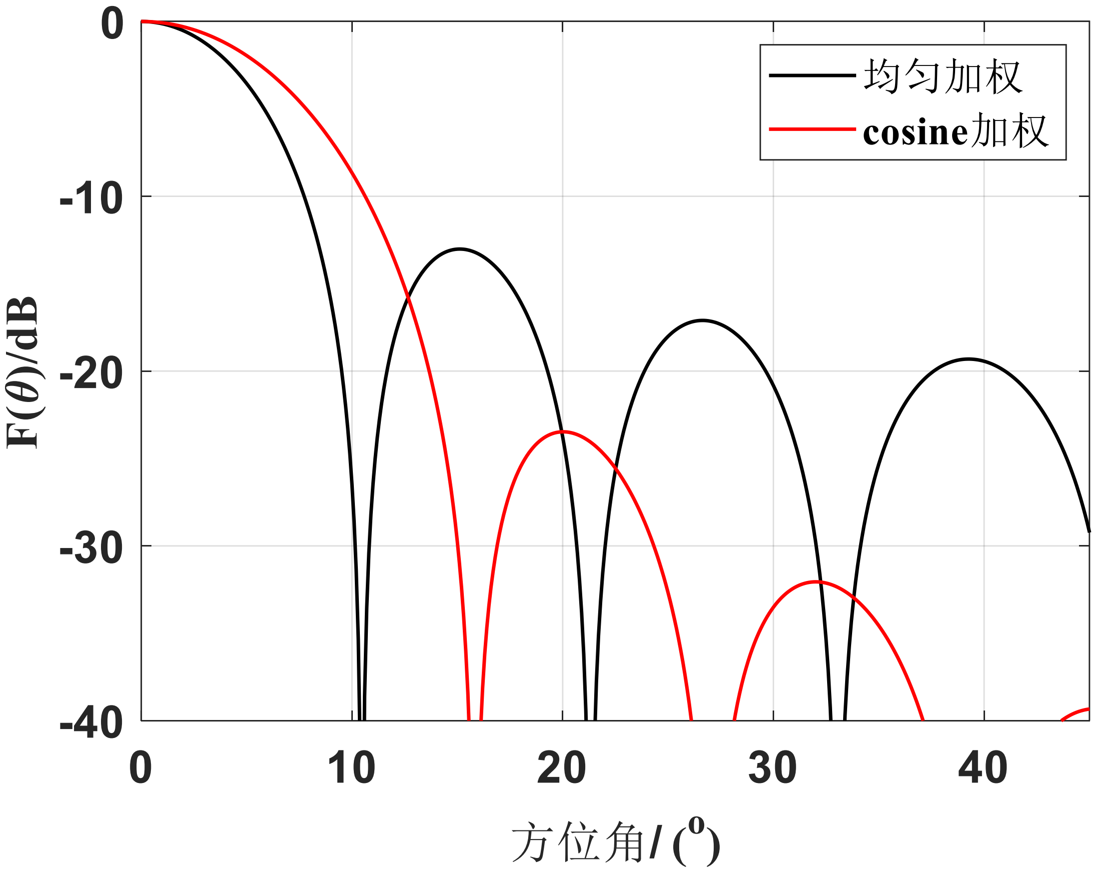
几种典型窗函数：

1. cosine加权

$$w_m = \sin\left(\frac{\pi}{2M}\right)\cos\left(\frac{\pi m}{M}\right),$$

$$-(M-1)/2 \le m \le (M-1)/2$$

相较于均匀加权，cosine加权拥有更低旁瓣，但其主瓣波束宽度也增加了。



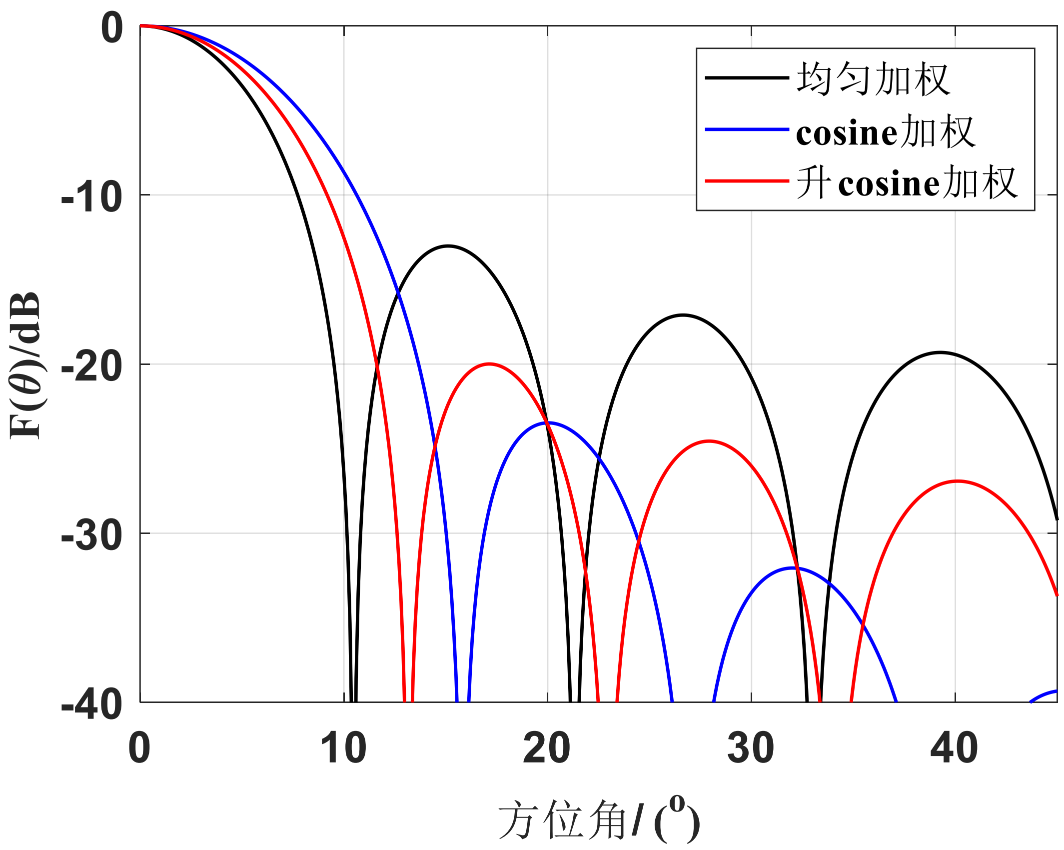
$M = 11$

2. 升cosine加权

$$w_m = c(p)\left[p + (1-p)\cos(\pi m/M)\right]$$

$$c(p) = p/M + (1-p)\sin(\pi/2M)/2$$

将矩阵加权和cosine加权结合起来，较均匀加权副瓣更低，较cosine加权主瓣展宽更小。

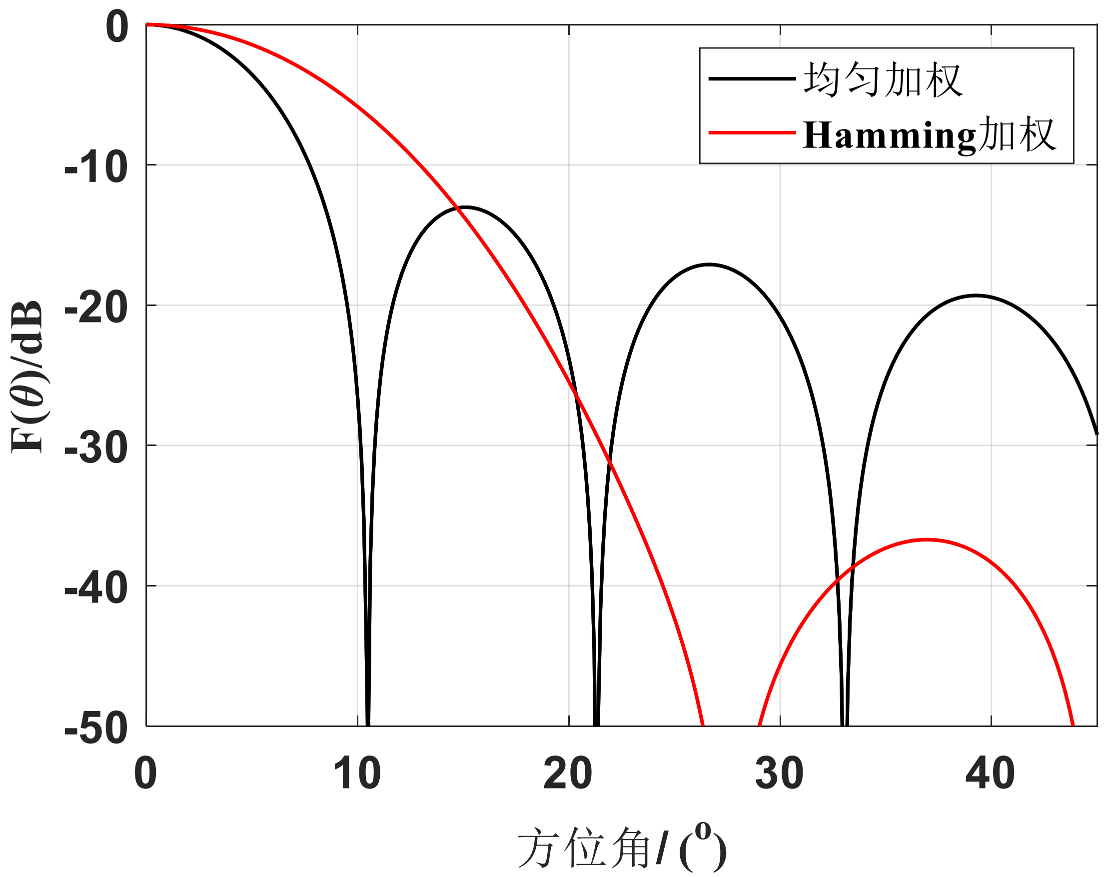


***M* = 11**

3. Hamming加权（改进升余弦窗）

$$w_m = g_0 + g_1 \cos\left(\frac{2\pi m}{M}\right)$$

利用矩形方向图和cosine平方方向图的特点，旁瓣更低，主瓣更宽。

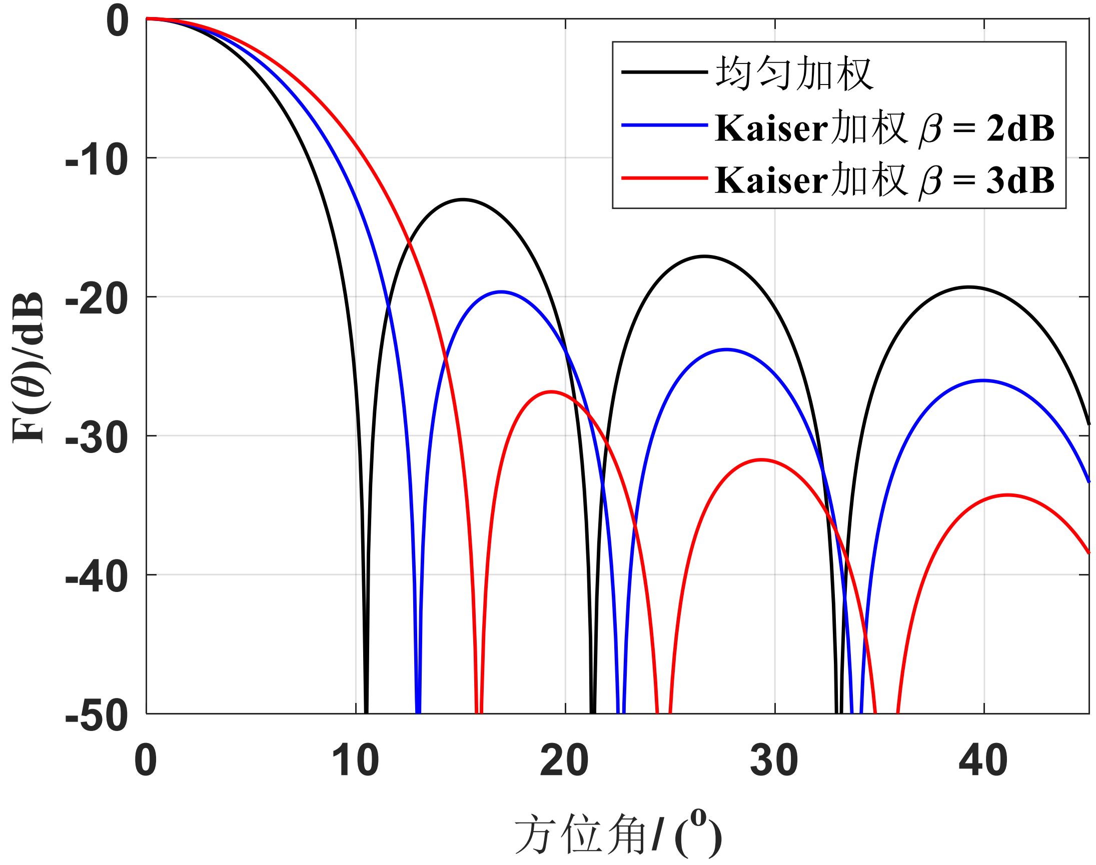


$M = 11$

4. Kaiser加权

$$w_m = I_0\left(\beta\sqrt{1-\left(\frac{2m}{M}\right)^2}\right)$$

I_0 是零阶修正Bessel函数，参数 β 决定了波束方向图在旁瓣峰值高度和主波束宽度之间的性能折中情况。

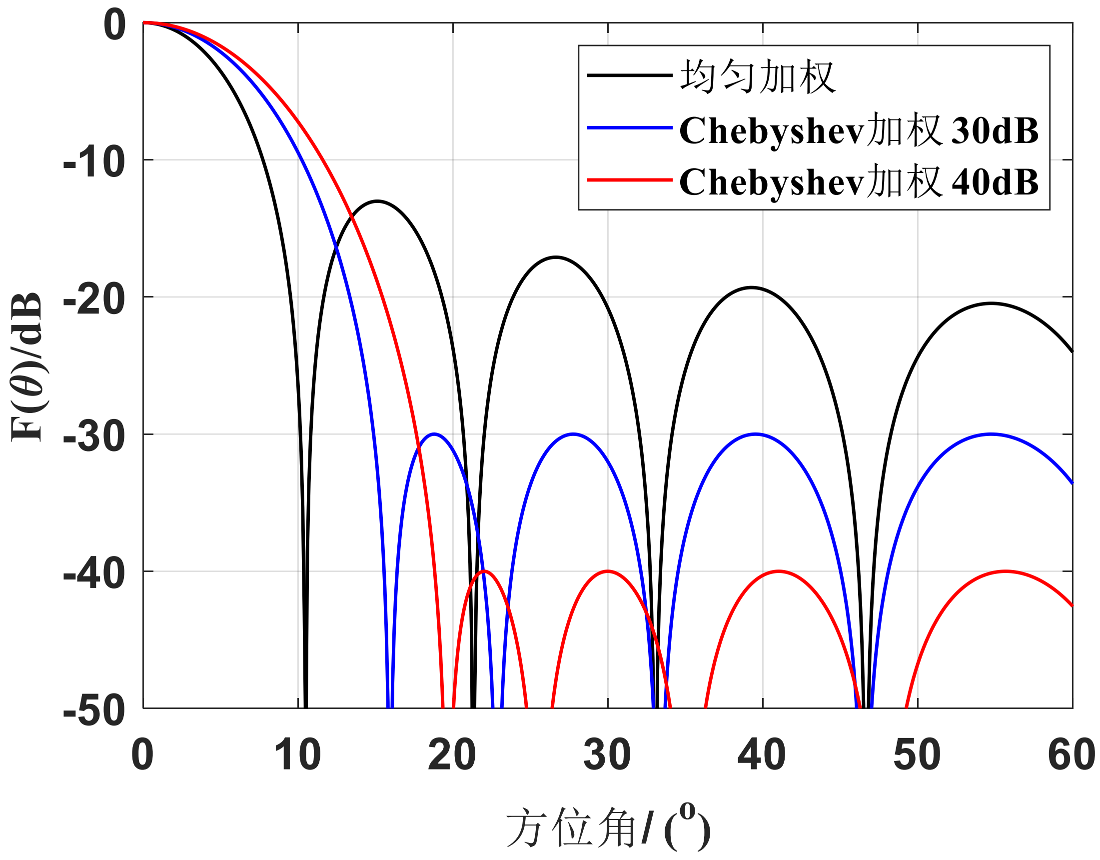


$M = 11$

5. Chebyshev加权

特点：

- 等副瓣电平；
- 在相同副瓣电平和相同阵列长度下主瓣最窄，为最佳阵列；
- 单元数过多时，阵列两端单元激励幅度跳变大，使馈电困难。

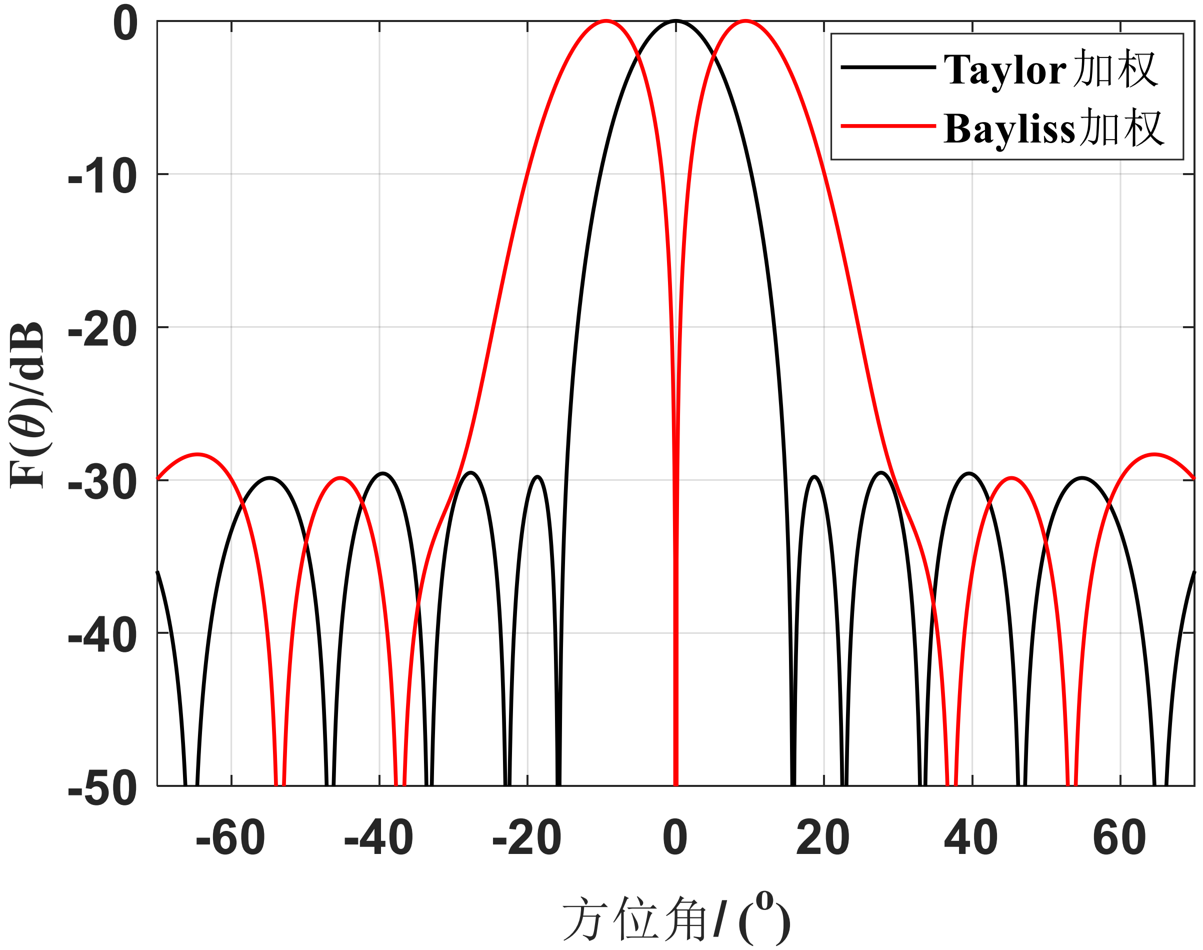


$M = 11$

6. Taylor加权和Bayliss加权

特点：

- **Taylor形成和波束；**
- **Bayliss形成差波束。**



$M = 11$

本次课内容

1. ADBF基本概念
2. 常规波束形成
3. 自适应波束形成

常规波束形成的优缺点：

优点： 是匹配滤波器，在主瓣方向信号相干积累，实现简单，在白噪声背景下是最优的，在而色噪声背景下，维纳滤波是最优的。

缺点：

- 波束宽度限制了方向角的分辨；
- 存在高旁瓣，强干扰可以从旁瓣进入；
- 加窗处理可降低旁瓣，但同时会展宽主瓣。

总之，常规波束形成依赖于阵列几何结构和波达方向角，而与信号环境无关，且固定不变，抑制干扰能力差。

自适应波束形成

将维纳滤波理论应用于空域滤波中，其权矢量依赖于信号环境。

一般框架：波束形成 $\mathbf{y}(n)=\mathbf{w}^{\mathrm{H}}\mathbf{x}(n)$

对于平稳随机信号，输出信号功率为：

$$E\left\{\left|\mathbf{y}(n)\right|^{2}\right\}=E\left\{\left[\mathbf{w}^{\mathrm{H}}\mathbf{x}(n)\right]\left[\mathbf{w}^{\mathrm{H}}\mathbf{x}(n)\right]^{\mathrm{H}}\right\}$$

$$=\mathbf{w}^{\mathrm{H}}E\left\{\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)\right\}\mathbf{w}$$

定义：阵列信号相关矩阵 $\mathbf{R}=E\left\{\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)\right\}$

包含了阵列接收信号的所有二阶统计知识。

统计最优波束形成 与 自适应波束形成

- 自适应阵列的权矢量根据阵列接收信号的统计特性求得，一般是该**二阶统计特性准确已知**情况下来分析，称为**统计最优波束形成**，相应地阵列权矢量为最优权矢量。

- 但是，阵列信号的统计特性一般是根据**有限快拍数据估计**得到，因此无法得到最优权矢量。有限快拍下的波束形成称为**自适应波束形成**，对应的阵列权矢量称为自适应权矢量。

$$\mathbf{R} = E\left\{\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)\right\} \Rightarrow \hat{\mathbf{R}} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)$$

$$\mathbf{R} = E\left\{\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)\right\}$$

$$= \begin{bmatrix} r_{11} & \cdots & r_{1M} \\ \vdots & \ddots & \vdots \\ r_{M1} & \cdots & r_{MM} \end{bmatrix}$$

$$\Rightarrow \quad \hat{\mathbf{R}} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)$$

$$= \begin{bmatrix} E\left\{\left|x_1(n)\right|^2\right\} & \cdots & E\left\{\left|x_1(n)x_M^*(n)\right|\right\} \\ \vdots & \ddots & \vdots \\ E\left\{\left|x_M(n)x_1^*(n)\right|\right\} & \cdots & E\left\{\left|x_M(n)\right|^2\right\} \end{bmatrix}$$

最优波束形成的寻优准则

- 最大信干噪比（MSINR）准则
- 最小均方误差（MMSE）准则
- 最小噪声方差（MNV）或线性约束最小方差（LCMV）准则

统计最优波束形成是一种分析工具，为自适应波束形成的实现提供了理论依据。

1. 最大信干噪比（MSINR）准则

阵列接收信号为：$\mathbf{x}(n)=\mathbf{x}_{\mathrm{S}}(n)+\mathbf{x}_{\mathrm{I+N}}(n)=\mathbf{a}(\theta_0)s(n)+\mathbf{x}_{\mathrm{I+N}}(n)$

目标信号协方差矩阵：

$$\begin{aligned}\mathbf{R}_{\mathrm{S}}&=E\left[\mathbf{a}(\theta_0)s(n)s^*(n)\mathbf{a}^{\mathrm{H}}(\theta_0)\right]\\&=\sigma_{\mathrm{S}}^2\mathbf{a}(\theta_0)\mathbf{a}^{\mathrm{H}}(\theta_0)\end{aligned}$$

干扰+噪声协方差矩阵：$\mathbf{R}_{\mathrm{I+N}}=E\left[\mathbf{x}_{\mathrm{I+N}}\mathbf{x}_{\mathrm{I+N}}^{\mathrm{H}}\right]$

MSINR的目的是使阵列输出信干噪比最大，即

$$\max_{\mathbf{w}}\mathrm{SINR}=\max_{\mathbf{w}}\frac{\mathbf{w}^{\mathrm{H}}\mathbf{R}_{\mathrm{S}}\mathbf{w}}{\mathbf{w}^{\mathrm{H}}\mathbf{R}_{\mathrm{I+N}}\mathbf{w}}$$

$$\max_{\mathbf{w}} \mathrm{SINR} = \max_{\mathbf{w}} \frac{\mathbf{w}^{\mathrm{H}}\mathbf{R}_{\mathrm{S}}\mathbf{w}}{\mathbf{w}^{\mathrm{H}}\mathbf{R}_{\mathrm{I+N}}\mathbf{w}}$$

将分母归一化，优化问题转化为

$$\begin{cases} \max\limits_{\mathbf{w}} \mathbf{w}^{\mathrm{H}}\mathbf{R}_{\mathrm{S}}\mathbf{w} \\ s.t.\ \ \mathbf{w}^{\mathrm{H}}\mathbf{R}_{\mathrm{I+N}}\mathbf{w} = 1 \end{cases}$$

利用拉格朗日乘子法，目标函数为 $L(\mathbf{w}) = \mathbf{w}^{\mathrm{H}}\mathbf{R}_{\mathrm{S}}\mathbf{w} + \lambda\left(\mathbf{I} - \mathbf{w}^{\mathrm{H}}\mathbf{R}_{\mathrm{I+N}}\mathbf{w}\right)$

求导等于零，可得 $\mathbf{R}_{\mathrm{S}}\mathbf{w} = \lambda\mathbf{R}_{\mathrm{I+N}}\mathbf{w}$

$$\mathbf{R}_S\mathbf{w} = \lambda\mathbf{R}_{I+N}\mathbf{w}$$

两边左侧各乘 $\mathbf{w}^H$，得到 $\mathbf{w}^H\mathbf{R}_S\mathbf{w} = \lambda\mathbf{w}^H\mathbf{R}_{I+N}\mathbf{w}$

可得 $\lambda = \dfrac{\mathbf{w}^H\mathbf{R}_S\mathbf{w}}{\mathbf{w}^H\mathbf{R}_{I+N}\mathbf{w}} = \mathrm{SINR}$

即特征值本身就是SINR，因此最优权矢量 $\mathbf{w}_{opt}$ 是与最大特征值对应的特征向量。

$$\mathbf{R}_S\mathbf{w} = \lambda\mathbf{R}_{I+N}\mathbf{w} \Rightarrow \mathbf{R}_{I+N}^{-1}\mathbf{R}_S\mathbf{w} = \lambda\mathbf{w}$$

即最优权矢量对应矩阵 $\mathbf{R}_{I+N}^{-1}\mathbf{R}_S$ 特征分解后最大特征值对应特征矢量。

$$\mathbf{R}_{\mathrm{S}}\mathbf{w} = \lambda \mathbf{R}_{\mathrm{I+N}}\mathbf{w}$$

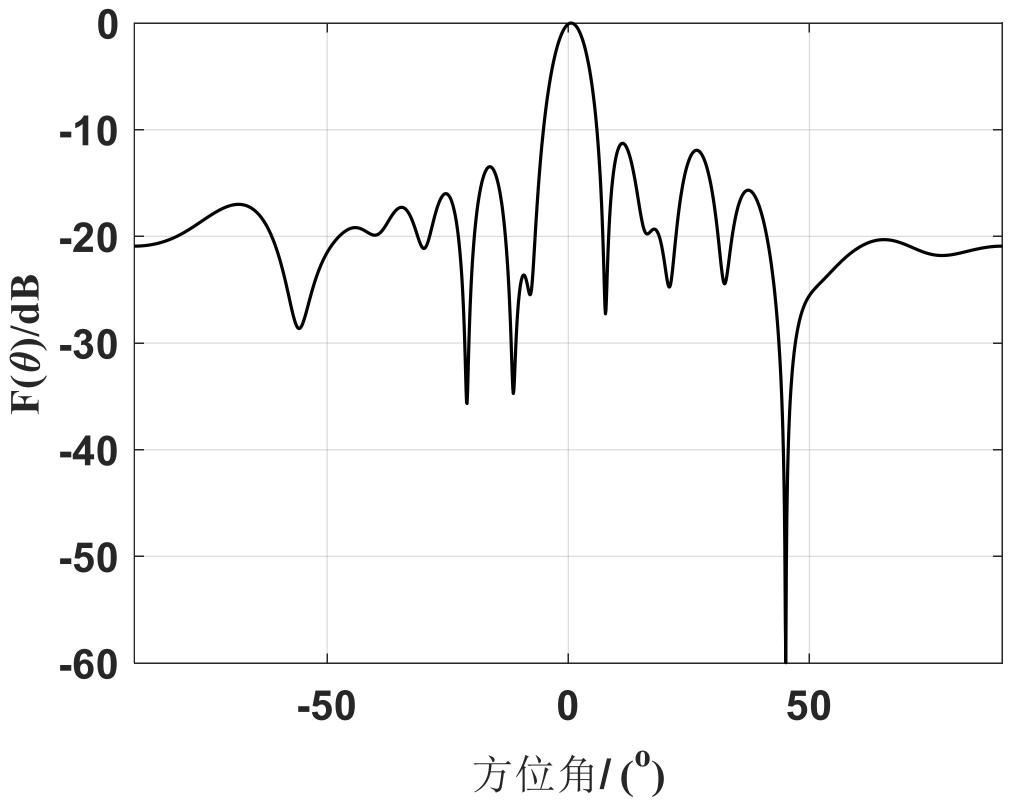
取最大特征值 $\lambda_{\max}$ 可得：

$$\begin{aligned}
&\mathbf{R}_{\mathrm{S}}\mathbf{w}_{\mathrm{opt}} = \lambda_{\max}\mathbf{R}_{\mathrm{I+N}}\mathbf{w}_{\mathrm{opt}} \\
\Rightarrow\ &\sigma_{\mathrm{S}}^{2}\mathbf{a}(\theta_0)\mathbf{a}^{\mathrm{H}}(\theta_0)\mathbf{w}_{\mathrm{opt}} = \lambda_{\max}\mathbf{R}_{\mathrm{I+N}}\mathbf{w}_{\mathrm{opt}} \\
\Rightarrow\ &\sigma_{\mathrm{S}}^{2}\mathbf{a}(\theta_0)c = \lambda_{\max}\mathbf{R}_{\mathrm{I+N}}\mathbf{w}_{\mathrm{opt}} \\
\Rightarrow\ &\mathbf{w}_{\mathrm{opt}} = \mu\mathbf{R}_{\mathrm{I+N}}^{-1}\mathbf{a}(\theta_0)
\end{aligned}$$

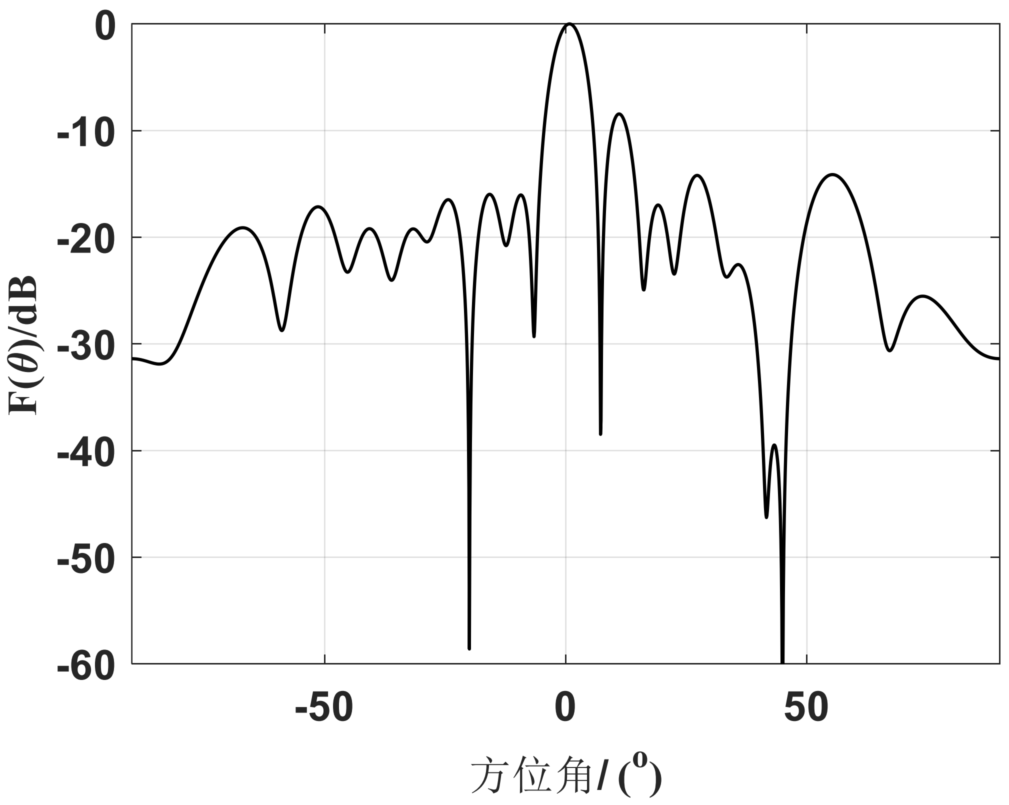
因此可得 MSINR准则下最优权矢量 $\mathbf{w}_{\mathrm{opt\text{-}MSINR}} = \mu\mathbf{R}_{\mathrm{I+N}}^{-1}\mathbf{a}(\theta_0)$

MSINR 准则自适应方向图

主波束方向（目标方向）： 0º



干扰方向：45º



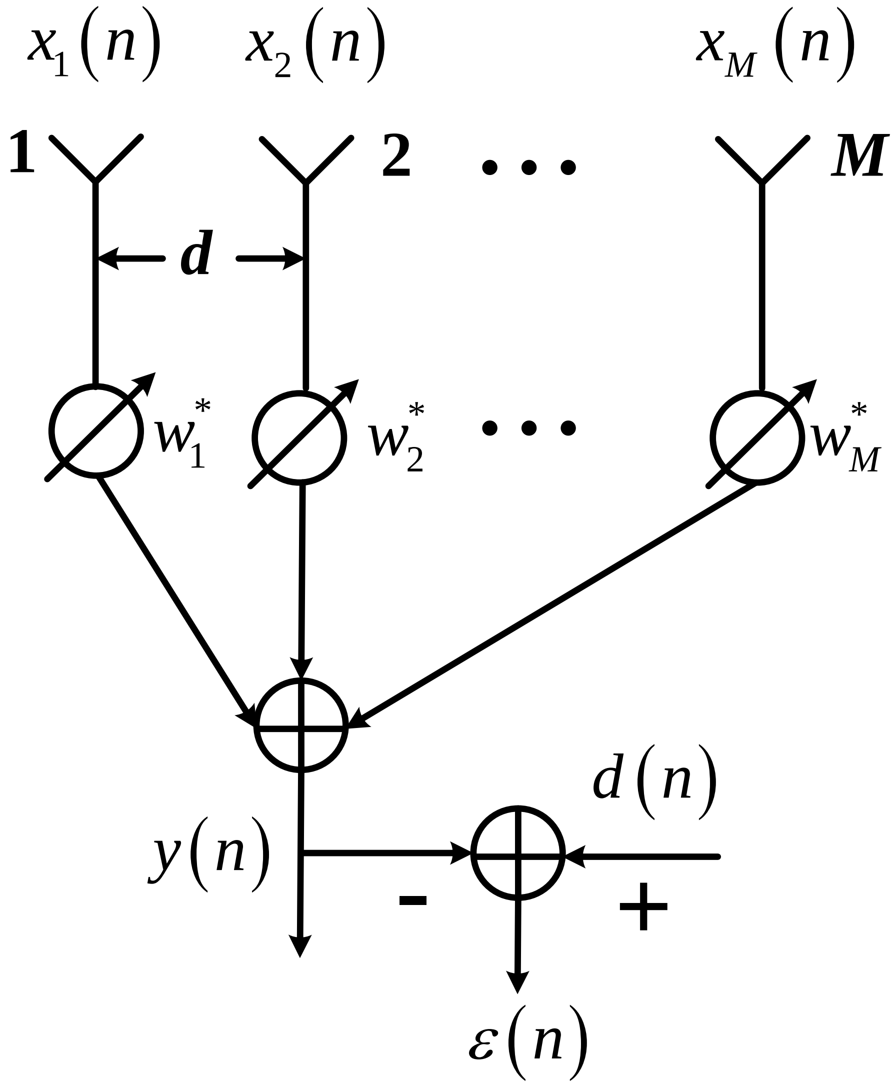
干扰方向：[-20º 45º]

2. 最小均方误差（MMSE）准则

- **利用参考信号求解自适应权矢量。**

  参考信号可根据期望信号产生的本地信号，也可以是接收的导引信号（如通信系统中导频信号）；

- **最优自适应权矢量的求取是使得参考信号与加权相加的阵列输出之差的均方值最小。**



阵列输出信号： $y(n)=\mathbf{w}^{\mathrm{H}}\mathbf{x}(n)$

令: $\sigma(\mathbf{w})=E\left[\left|y(n)-d(n)\right|^2\right]=E\left[\left|\mathbf{w}^{\mathrm{H}}\mathbf{x}(n)-d(n)\right|^2\right]$

则目标为: $\min_{\mathbf{w}}\sigma(\mathbf{w})$

$$\sigma(\mathbf{w})=E\left[\left|\mathbf{w}^{H}\mathbf{x}(n)-d(n)\right|^2\right]=E\left\{\left[\mathbf{w}^{\mathrm{H}}\mathbf{x}(n)-d(n)\right]\left[\mathbf{w}^{\mathrm{H}}\mathbf{x}(n)-d(n)\right]^{\mathrm{H}}\right\}$$

$$=\mathbf{w}^{\mathrm{H}}\mathbf{R}_x\mathbf{w}+E\left[\left|d(n)\right|^2\right]-\mathbf{w}^{\mathrm{H}}\mathbf{r}_{xd}-\mathbf{r}_{xd}^{\mathrm{H}}\mathbf{w}$$

其中 $\mathbf{r}_{xd}=E\left[\mathbf{x}(n)d^*(n)\right]$ **是互相关矢量**

$\mathbf{R}_x=E\left[\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)\right]$ **是自相关矩阵**

对 w 求导并令其为 0，可得：$\mathbf{w}_{\text{opt-MMSE}} = \mathbf{R}_{\text{x}}^{-1}\mathbf{r}_{xd}$

该解是最优维纳解形式，更多被应用于旁瓣相消处理。

由公式可看出：应用此方法仅需阵列接收信号的自相关矩阵和接收信号与参考信号的互相关矢量。

MMSE准则的应用：

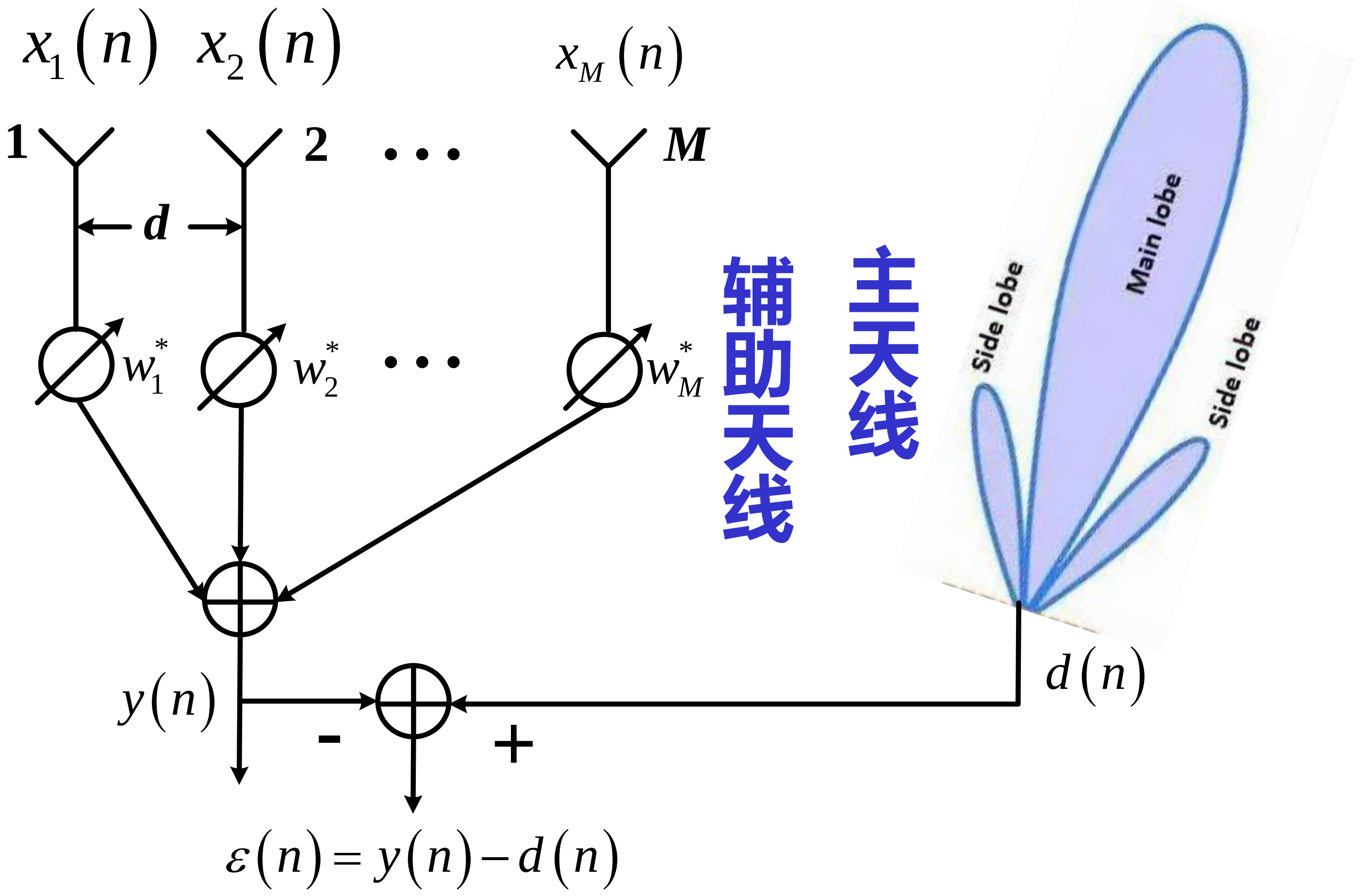
- 自适应均衡（通信）
- 多通道均衡（雷达）
- 自适应天线旁瓣相消（SLC）

实例：天线旁瓣相消技术 （Antenna Sidelobe Canceler ）

辅助天线：

增益小，选取与主天线旁瓣电平相当，无方向性，因此 y(n) 仅为干扰信号。

权矢量： $\mathbf{w}_{\text{opt-MMSE}} = \mathbf{R}_{\text{x}}^{-1}\mathbf{r}_{xd}$

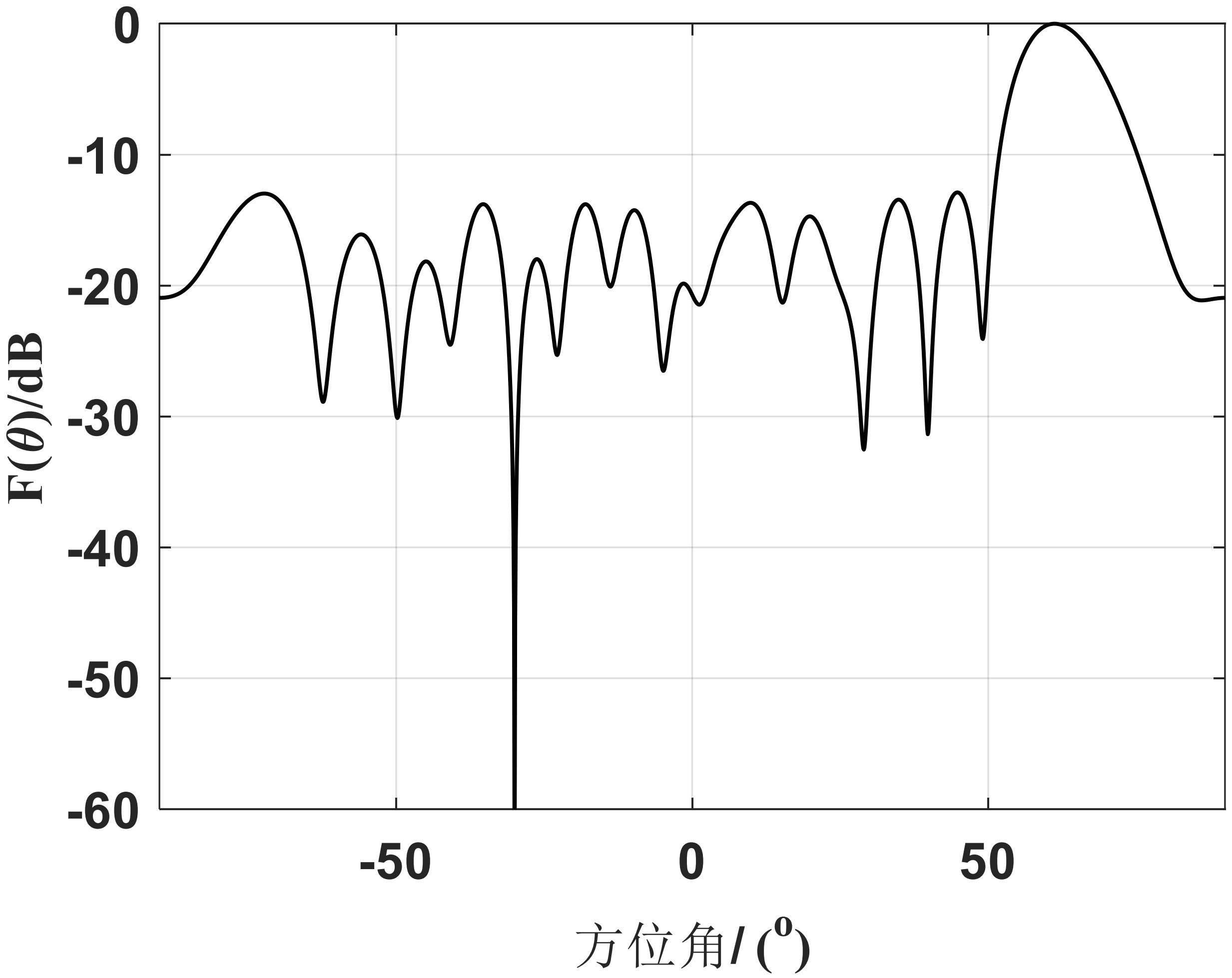


良好干扰抑制性能条件：主天线与辅助天线对接收干扰信号相关性较好。

MMSE 准则自适应方向图

目标方向：60°

干扰方向：-30°



3. 线性约束最小方差（LCMV）准则

主要思想：期望信号响应为常数前提下阵列输出功率最小。

阵列输出： $y(n)=\mathbf{w}^{\mathrm{H}}\mathbf{x}(n)$

方差： $E\left[\left|y(n)\right|^{2}\right]=\mathbf{w}^{\mathrm{H}}\mathbf{R}\mathbf{w}$ **即输出功率**

导向矢量约束$\boldsymbol{a}(\boldsymbol{\theta}_{\mathbf{0}})$为目标信号导向矢量。

优化函数为： $\begin{cases} \min\limits_{\mathbf{w}} \mathbf{w}^{\mathrm{H}}\mathbf{R}_x\mathbf{w} \\ s.t.\ \mathbf{w}^{\mathrm{H}}\mathbf{a}(\theta_0)=1 \end{cases}$ **1 可变为任意非零常数。**

解得： $\mathbf{w}_{\text{opt-LCMV}} = \mu \mathbf{R}_x^{-1}\mathbf{a}(\theta_0)$ μ **可变为任意非零常数。**

如果固定： $\mathbf{w}^{\mathrm{H}}\mathbf{a}(\theta_0)=1$ **则** $\mu = \dfrac{1}{\mathbf{a}^{\mathrm{H}}(\theta_0)\mathbf{R}_x^{-1}\mathbf{a}(\theta_0)} = 1$

式中 μ 的取值不影响 SNR 和方向图。

注：本准则要求波束形成的指向$a(\theta_0)$已知，但不要求参考信号 d(n)和信号与干扰的互相关矢量。

推广到约束多个方向：一般的线性约束最小方差法为

$$\begin{cases} \min\limits_{\mathbf{w}} \mathbf{w}^{\mathrm{H}} \mathbf{R}_x \mathbf{w} \\ s.t.\ \mathbf{w}^{\mathrm{H}} \mathbf{C} = \boldsymbol{\beta}^{\mathrm{H}} \end{cases} \qquad \mathbf{C} \in \mathbb{C}^{M \times L}, \boldsymbol{\beta} \in \mathbb{R}^{L \times 1}$$

解： $\mathbf{w}_{\text{opt-LCMV}} = \mathbf{R}_x^{-1} \mathbf{C} \left(\mathbf{C}^{\mathrm{H}} \mathbf{R}_x^{-1} \mathbf{C} \right)^{-1} \boldsymbol{\beta}$

特例：当 $\mathbf{C} = \mathbf{a}(\theta_0)$**，即约束单个方向，则** $\boldsymbol{\beta} = 1$**。**

实际应用：

当已知目标在 θ_0 方向，但也可能在 θ_0 附近，这时可令

$$\begin{cases} \mathbf{C} = \left[\mathbf{a}(\theta_0), \mathbf{a}(\theta_0 + \Delta\theta), \mathbf{a}(\theta_0 - \Delta\theta)\right] \\ \beta = [1,1,1]^{\mathrm{T}} \end{cases}$$

可增加算法稳健性。

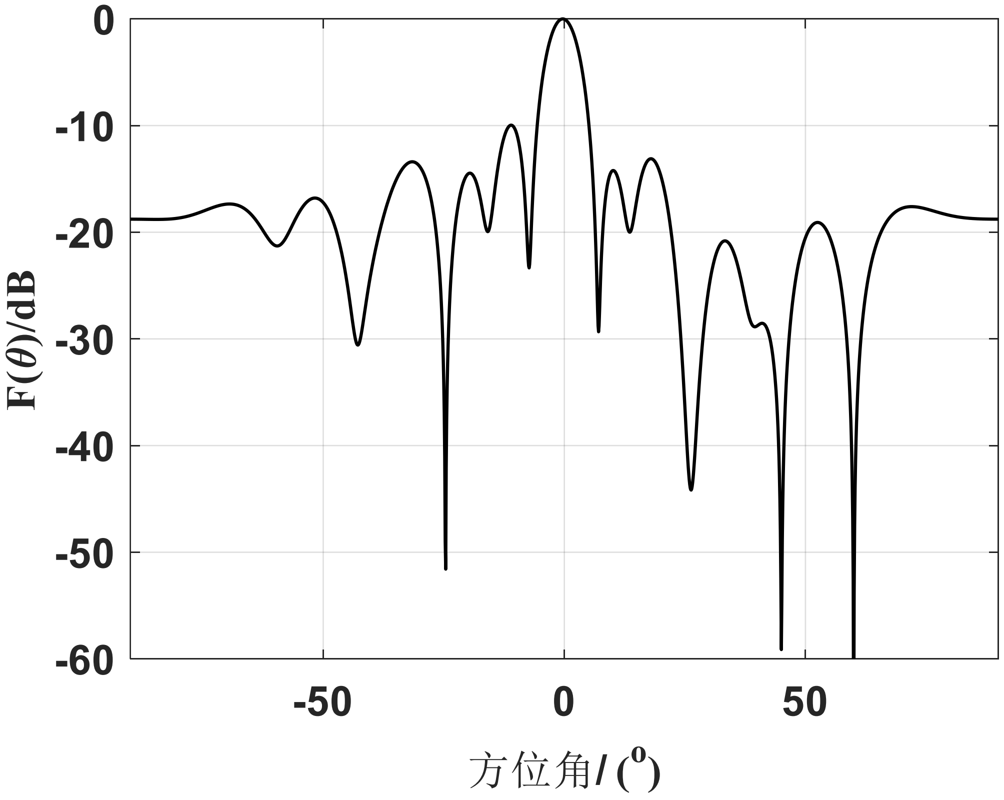
注： 针对白噪声，$\mathbf{R}_x$ 为单位阵，阵列输出为 $\mathbf{w}^H\mathbf{R}_x\mathbf{w} = \mathbf{w}^H\mathbf{w}$，此时自适应滤波无能为力。换句话说，自适应滤波无法滤除白噪声。

- 在实际工程中，阵列天线不可避免地存在各种误差。
- 文献 Error analysis of the optimal antenna array processors. *IEEE Trans. On AES*, 1986, 22(3):395-409. 对各种误差（如阵元误差、通道频率响应误差、阵元位置扰动误差、互耦等）的影响进行了综述。

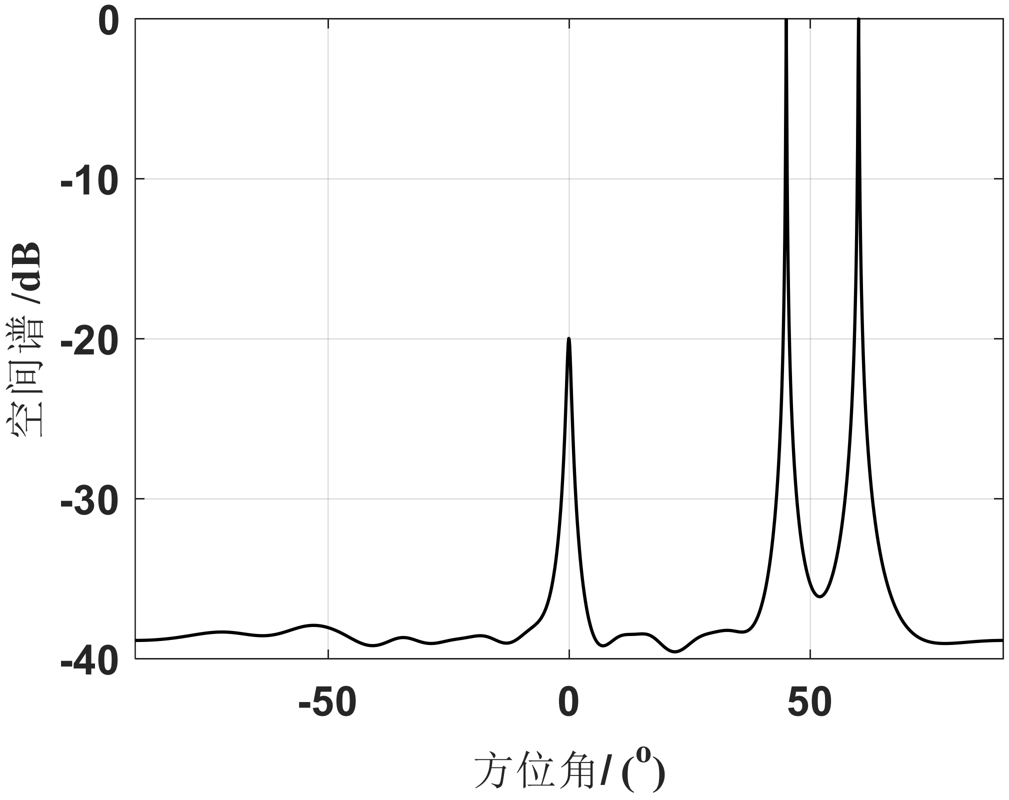
基本结论：

- 对于只利用干扰 + 噪声协方差矩阵求逆的自适应方法，幅相误差对自适应波束形成的影响不大；
- 对于利用信号 + 干扰 + 噪声协方差矩阵求逆的自适应方法，当信噪比较大时，虽然干扰零点位置变化不大，但是在信号方向上也可能形成零陷，导致信噪比严重下降。

LCMV准则，对应Capon方向图和Capon谱 目标方向：0^o 干扰方向：$[45^o\ 60^o]$



Capon自适应方向图

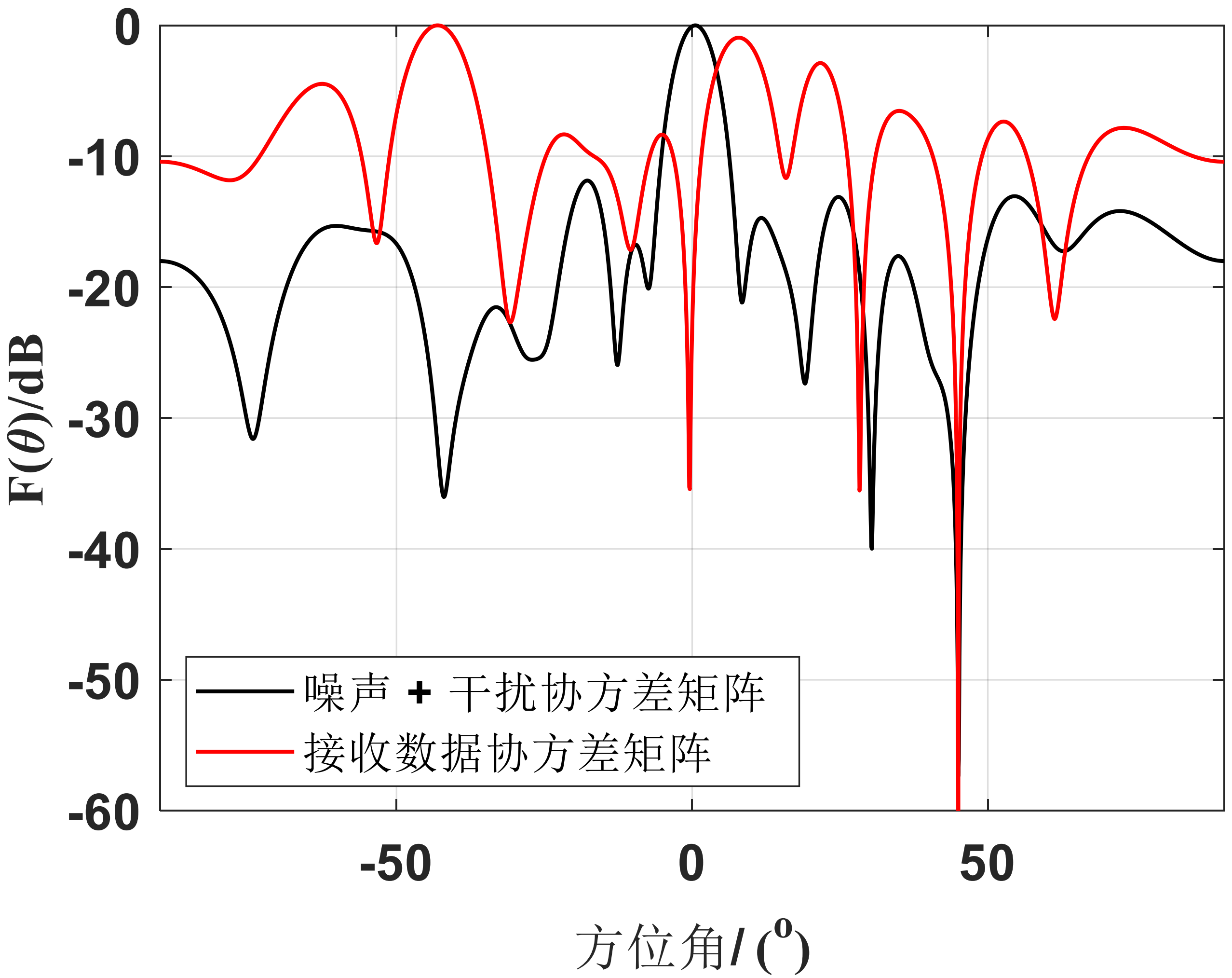


Capon功率谱

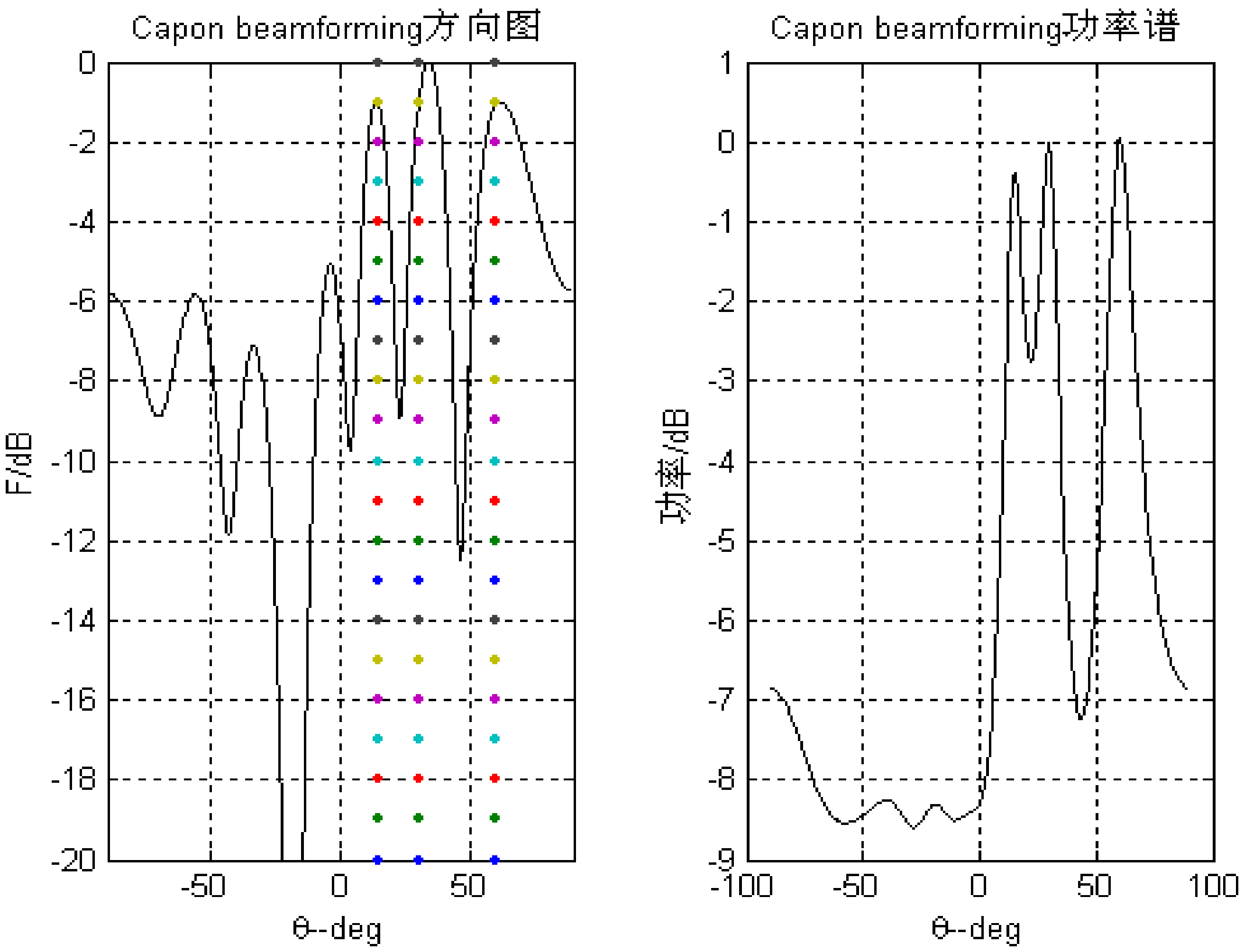
Capon法自适应方向图

$$\mathbf{w}_{\text{opt-LCMV}} = \mu \mathbf{R}_x^{-1} \mathbf{a}\left(\theta_0\right)$$

- **当协方差矩阵含有目标信号时，发生目标相消现象，导致目标能量损失而无法被检测；**
- **应采用干扰 + 噪声协方差矩阵形成自适应权值。**



多约束波束形成，来波方向 0 20 50



4. 三个最优准则的内在关系

准则	解的表达式	所需已知条件
MSINR	$\mathbf{w}_{\text{opt-MSINR}} = \mu \mathbf{R}_{\text{I+N}}^{-1}\mathbf{a}(\theta_0)$	已知 $\mathbf{R}_{\mathbf{I+N}}$
MMSE	$\mathbf{w}_{\text{opt-MMSE}} = \mathbf{R}_{\text{x}}^{-1}\mathbf{r}_{xd}$	已知期望信号d(n)
LCMV	$\mathbf{w}_{\text{opt-LCMV}} = \mu \mathbf{R}_{x}^{-1}\mathbf{a}(\theta_0)$	已知期望信号方向θ_0

MSINR 和 LCMV 准则的等效性

MSINR准则下最优权矢量： $\mathbf{w}_{\text{opt-MSINR}} = \mu \mathbf{R}_{\text{I+N}}^{-1} \mathbf{a}(\theta_0)$

LCMV 准则下最优权矢量： $\mathbf{w}_{\text{opt-LCMV}} = \mu \mathbf{R}_{x}^{-1} \mathbf{a}(\theta_0)$

注意： $\mathrm{R_x}$ 中含有期望信号分量，而 $\mathrm{R_{I+N}}$ 中不含期望信号分量，仅为噪声分量。

即：

$$\mathbf{R}_x = \mathbf{R}_s + \mathbf{R}_{\text{I+N}} = \sigma_s^2 \mathbf{a}(\theta_0) \mathbf{a}^{\text{H}}(\theta_0) + \mathbf{R}_{\text{I+N}}$$

由矩阵求逆引理： $\left(\mathbf{b}\mathbf{b}^{\mathrm{H}}+\mathbf{A}\right)^{-1}=\mathbf{A}^{-1}-\dfrac{\mathbf{A}^{-1}\mathbf{b}\mathbf{b}^{\mathrm{H}}\mathbf{A}^{-1}}{1+\mathbf{b}^{\mathrm{H}}\mathbf{A}^{-1}\mathbf{b}}$ **假设 A 可逆**

所以： $\mathbf{R}_x^{-1}=\mathbf{R}_{\mathrm{I+N}}^{-1}-\dfrac{\sigma_s^2\mathbf{R}_{\mathrm{I+N}}^{-1}\mathbf{a}(\theta_0)\mathbf{a}^{\mathrm{H}}(\theta_0)\mathbf{R}_{\mathrm{I+N}}^{-1}}{1+\sigma_s^2\mathbf{a}^{\mathrm{H}}(\theta_0)\mathbf{R}_{\mathrm{I+N}}^{-1}\mathbf{a}(\theta_0)}$

可得：

$$\begin{aligned}\mathbf{w}_{\text{opt-LCMV}}&=\mu\mathbf{R}_x^{-1}\mathbf{a}(\theta_0)\\&=\mu\mathbf{R}_{\mathrm{I+N}}^{-1}\mathbf{a}(\theta_0)\left[1-\frac{\sigma_s^2\mathbf{R}_{\mathrm{I+N}}^{-1}\mathbf{a}(\theta_0)\mathbf{a}^{\mathrm{H}}(\theta_0)}{1+\sigma_s^2\mathbf{a}^{\mathrm{H}}(\theta_0)\mathbf{R}_{\mathrm{I+N}}^{-1}\mathbf{a}(\theta_0)}\right]\\&=\bar{\mu}\mathbf{R}_{\mathrm{I+N}}^{-1}\mathbf{a}(\theta_0)\quad\rightarrow\quad\mathbf{w}_{\text{opt-MSINR}}\end{aligned}$$

上式表明： 在精确的方向矢量约束条件和相关矩阵精确已知，且目标功率较小， MSINR 准则与 LCMV 准则等效。

如不满足： 应用 $\mathbf{R}_{\mathbf{I+N}}$ 来计算权值，直接用 $\mathbf{R}_{\mathbf{x}}$ 计算最优权值会导致信号相消。

最优波束形成方法中，降低旁瓣电平的方法是加窗处理。

$\mathbf{a}_h(\theta_0)=\mathbf{h}\cdot\mathbf{a}(\theta_0)$ **h** 为加窗矩阵。

$$\begin{cases}\min\limits_{\mathbf{w}} \mathbf{w}^{\mathrm{H}}\mathbf{R}_x\mathbf{w} \\ s.t.\ \mathbf{w}^{\mathrm{H}}\mathbf{a}_h(\theta_0)=1\end{cases} \Rightarrow \mathbf{w}_{\mathrm{opt}}=\mu\mathbf{R}_x^{-1}\mathbf{a}_h(\theta_0)$$

MMSE 准则与MSINR 和 LCMV 准则的等效性

MMSE 准则下最优权矢量： $\mathbf{w}_{\text{opt-MMSE}} = \mathbf{R}_{\text{x}}^{-1}\mathbf{r}_{xd}$

若已知 $\mathbf{x}_{\text{I+N}}(n)$ 与 $\mathbf{d}(n)$ 不相关，则

$$\begin{aligned}\mathbf{r}_{xd} &= E\left[\mathbf{x}(n)d^*(n)\right]\\ &= E\left[s(n)\mathbf{a}(\theta_0)d^*(n)\right]\\ &= E\left[s(n)d^*(n)\right]\mathbf{a}(\theta_0)\end{aligned}$$

因此， $\mathbf{w}_{\text{opt-MMSE}} = E\left[s(n)d^*(n)\right]\mathbf{R}_{\text{x}}^{-1}\mathbf{a}(\theta_0) = \mu\mathbf{R}_{\text{x}}^{-1}\mathbf{a}(\theta_0)$ $\qquad \mu = E\left[s(n)d^*(n)\right]$

由此看出，上述三个准则在一定条件下是等价的。

自适应波束形成小结：

$$\mathbf{w}_{\text{opt}} = \mu \mathbf{R}_{\text{I+N}}^{-1} \mathbf{a}(\theta_0)$$

- 需已知二阶统计量 $\mathbf{R}_{\text{I+N}}$；
- 已知 $\mathbf{a}(\theta_0)$；
- 矩阵求逆运算量大，有待于寻找快速算法。