

一种不被重视的自适应波束形成的算法

E. B. Lunde

(挪威防御研究中心)

一、引言

有许多完全不同的算法可以用于自适应波束形成,但大部分文献处理的是梯度下降(GD)算法。梯度下降算法的特点、稳定性以及性能是大家所熟知的。本文强调 Woodbury 恒等(WI)算法。这个名字起因于该算法用了 Woodbury 恒等式(矩阵求逆的辅助定理)。用到 GD 算法时只是作为参考。

在最佳波束形成的许多准则中,本文所用的是有约束的最小均方(LMS)准则。它要求波束所指方向上频率响应无失真同时输出噪声功率最小。

这里将要讨论 WI 算法的数值稳定性以及确保这种稳定度的各种方法。然后比较两种算法对计算机的要求、性能和灵活性。

二、基 阵 处 理

研究图 2.1 所示 K 通道基阵处理器。假定其输入是一组均值为零、经过离散付氏变换的传感器信号,为方便起见,把它们写成数据输入矢量:

$$\mathbf{X}^T = [x_1, x_2, \dots, x_K] \quad (2.1)$$

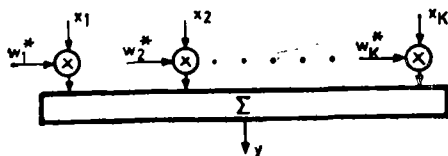


图2.1 基 阵 处 理 器

用同样的方法定义权重矢量

$$\mathbf{W}^T = [w_1, w_2, \dots, w_K] \quad (2.2)$$

则处理器的输出是

$$\mathbf{y} = \mathbf{W}^+ \cdot \mathbf{X} \quad (2.3)$$

*原文题目: "The Forgotten Algorithm in Adaptive Beamforming."

上角标 $*$, T , $^+$ 分别表示复共轭、转置和复共轭转置。

1. 最佳基阵处理

X 的期望互谱密度矩阵用

$$R = E\{X \cdot X^+\} \quad (2.4)$$

来定义。理想条件下, 空间点源 Q 将对 R 有贡献。此贡献用并矢^[1]给出:

$$R_Q = \sigma_Q^2 d \cdot d^+ \quad (2.5)$$

这里, σ_Q^2 是传感器上 Q 的功率谱密度, d 是 Q 相对于给定基阵几何位置的相位矢量。

把传感器阵聚焦到点 Q 时, 按上述准则得到的最佳权重矢量是

$$W_{opt}^+ = m^+ R^{-1} / (m^+ R^{-1} m) \quad (2.6)$$

这里, 聚焦矢量 m 必须选得等于点 Q 的相位矢量:

$$m = m(r, \phi) = d \quad (2.7)$$

通常选 m 使基阵聚焦到无穷远处, 因此 $m = m(\phi)$, 叫做指向矢量。处理器 (波束形成器) 有 L 个不同的指向矢量, 就说它有 L 个波束。

2. 自适应基阵处理

自适应算法的任务是在 R 并不精确已知的情况下, 给出不断修正的 W_{opt} 的估值。梯度下降算法和 Woodbury 算法解决这个课题的方法完全不同。梯度算法直接调整权重矢量^[2]:

$$y_l(n) = W_l^+(n) X(n) \quad l = 1, 2, \dots, L \quad (2.8)$$

$$W_l(n+1) = (I - m_l m_l^+ / K) [W_l(n) - \mu_{GD} y_l^*(n) \cdot X(n)] + m_l / K \quad l = 1, 2, \dots, L \quad (2.9)$$

其中 μ_{GD} 是确定自适应时间常数的参数, $0 < \mu_{GD} \leq 1$ 。

Woodbury 算法不直接调整权重 W , 而是调整矩阵 R^{-1} 和标量 b 。 b 的定义是

$$b = m^+ R^{-1} m \quad (2.10)$$

处理程序可以写作^[3]

$$D(n) = R^{-1}(n) \cdot X(n) \quad (2.11)$$

$$h(n) = [X^+(n) \cdot D(n) + (1 - \mu_{WI}) / \mu_{WI}]^{-1} \quad (2.12)$$

$$R^{-1}(n+1) = [R^{-1}(n) - h(n) \cdot D(n) \cdot D^+(n)] / (1 - \mu_{WI}) \quad (2.13)$$

$$g_l(n) = m_l^+ D(n) \quad l = 1, 2, \dots, L \quad (2.14)$$

$$y_l(n) = g_l(n) / b_l(n) \quad l = 1, 2, \dots, L \quad (2.15)$$

$$b_l(n+1) = [b(n) - h(n) \cdot g_l^*(n) \cdot g_l(n)] / (1 - \mu_{WI}) \quad l = 1, 2, \dots, L \quad (2.16)$$

为得到(2.13), 用到了 Woodbury 恒等式, 因此这个算法的名字叫做 Woodbury 算法。注意, 对所有波束所指的方向来说, (2.13) 式是共用的。在比较两算法以前, 先讨论 WI 算法的稳定性。

三、Woodbury 恒等算法的数值稳定性

可以证明, 在一定条件下, 调整过程中 R^{-1} 数值上是稳定的^[4]。若计算中所用字长足够长, 在实际情况中这个条件就满足。假定, 用整型矩阵 P 和整数 q 表示 R^{-1} , 即

$$R^{-1} = 2^q P + \text{舍入误差矩阵} \quad (3.1)$$

对矩阵 P 的元素来说, 足够的字长长度 B_1 是三个参数的函数:

$$2^{B_1} \sim c / (g \sqrt{\mu_{WI}}) \quad (3.2)$$

μ_{WI} 是决定调整过程时间常数的参数。 g 是最大相对输出误差，

$$g = \max_1 [\{|y_1(n)|^2 - |y_{1,opt}|^2\} / |y_{1,opt}|^2] \quad (3.3)$$

c 是矩阵 R (或 R^{-1}) 的条件数

$$c = \lambda_{\max} / \lambda_{\min} \quad (3.4)$$

其中 λ_i , $i=1, 2, \dots, K$ 是矩阵 R 的本征值。 B_1 也是矩阵维数 K 的隐函数，这是因为

$$c \leq \text{常数} \cdot K - 1 \quad (3.5)$$

μ_{WI} 和 g 都是已知量(选定量)，因此，唯一能控制来加强稳定性的一个量就是条件数 c 。若数据的状况是 $c > c_0$ ，这里 $c = c_0$ 用来计算 B_1 ，则该数据必须修改成 $c_{\text{mod}} \leq c_0$ 。最简单也是最自然的修改方法是在数据上加一人工白噪声：

$$X_{\text{mod}} = X + X_{\text{人工}} \quad (3.6)$$

人工白噪声的互谱矩阵是

$$R_{\text{人工}} = E\{X_{\text{人工}} X_{\text{人工}}^+\} = \sigma_{\text{人工}}^2 I \quad (3.7)$$

这就使条件数减少成

$$c_{\text{mod}} = (\lambda_{\max} + \sigma_{\text{人工}}^2) / (\lambda_{\min} + \sigma_{\text{人工}}^2) < c \quad \text{当 } \lambda_{\max} > \lambda_{\min} \quad (3.8)$$

若 $\lambda_{\max}$ 和 $\lambda_{\min}$ 已知，可用 $c_{\text{mod}} = c_0$ 求得所需 $\sigma_{\text{人工}}^2$ 的数值：

$$\sigma_{\text{人工}}^2 = \begin{cases} (\lambda_{\max}/c_0) - \lambda_{\min} & c > c_0 \\ 0 & c \leq c_0 \end{cases} \quad (3.9)$$

这个方法需要计算 $\lambda_{\max}$ 和 $\lambda_{\min}$ 。另一方法虽然要求的计算较少，但得到的性能也就较差，该方法使用公式

$$\sigma_{\text{人工}}^2 = \text{tr}(R) / c_0 \quad (3.10)$$

对给定的条件数 c ，这个方法得到的 c_{mod} 介于最大值和最小值之间，如图 3.1 所画的那样。图 3.1 同时画出第一个方法得到的 c_{mod} ，由此看出第一个方法的结果较好。

与 R^{-1} 的调整情况相反， $b = m^+ R^{-1} m$ 的调整在数值上是不稳定的。因此自适应 M 次以后，必须直接计算 b 。 M 是两个参数的函数^[4]：

$$M \sim \Delta B_1 / \mu_{WI} \quad (3.11)$$

这里 ΔB_1 是比特数。为稳定调正 R^{-1} ，字长 B_1 要比必须的大一些。因此，在字长 B_1 和需要在多大频繁程度上直接计算 b 二者之间往往要作折衷考虑。 b 的表达式中足够的字长 B_2 由

$$2^{B_2} \sim [\exp(\mu_{WI} M) - 1] / (\mu_{WI} g) \quad (3.12)$$

决定。

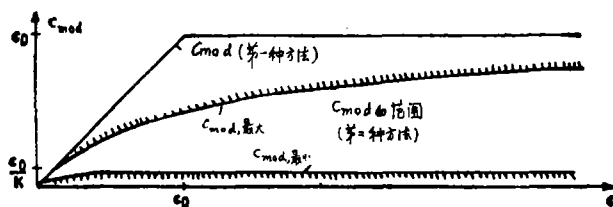


图 3.1 c_{mod} 的范围(作为 c 的函数)

四、两个算法的比较

在对计算机的要求、算法的性能和灵活性等方面来比较 GD 和 WI 两算法。

1. 对计算机的要求

从两个方面来评价两个算法对计算机的要求：实乘次数表示的计算能力和存储量的大小。不考虑字长的不同。据此，图 4.1、图 4.2 把实乘次数分别表示成传感器数和波束数的函数。图上的 CON 表示普通波束形成(延迟并相加)。图 4.3 表示两个算法实乘次数的比。根据经验方法，两个算法之间的过渡界($\beta=1$)可以按 $K \approx L$ ，即传感器数近似等于波束数来确定。

WI 算法对大 L 的情况较适用，因为对所有波束来讲， R^{-1} 的调正过程是一样的。

与此相似，图 4.4 和图 4.5 把所需存储量表示成传感器数和波束数的函数。图 4.6 示出的是两种算法对存储量要求的比。

这些图和曲线只是大致的特征的表示，但至少说明，一种算法并不总是胜过另一算法。

2. 性能

用满足两个相互矛盾的要求的能力来评价算法性能的好坏：既要求调整到尽可能没有噪声(长时间常数)，同时又要尽可能快地自适应于非平稳状态(短时间常数)。定量表示性能的相当好的手段是所谓的学习曲线。本文中用

$$P(n) = \left\{ \sum_{i=1}^L [W_i^+(n) R W_i(n) - W_{i,opt}^+ R W_{i,opt}] \right\} / L \quad (4.1)$$

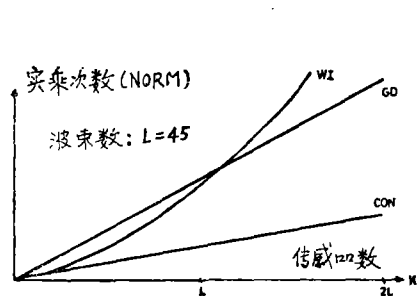


图 4.1 实乘次数作为传感器数的函数

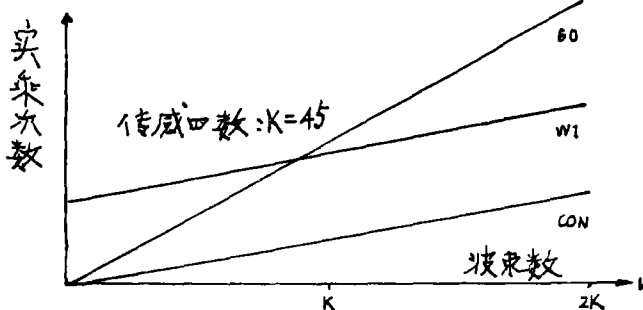


图 4.2 实乘次数作为波束数的函数

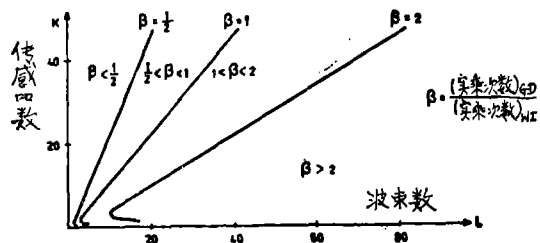


图 4.3 实乘次数的比(以传感器数和波束数为自变量)

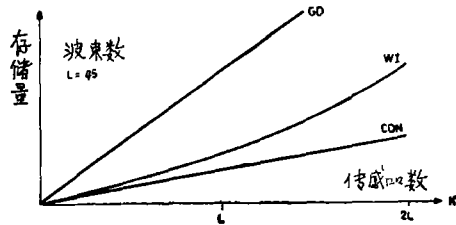


图 4.4 存储量作为传感器数的函数

定义学习曲线。这里， l 是表示波束的下标。梯度算法中 $W_l(n)$ 可直接得到，Woodbury 算法用

$$W_l^+(n) = m_l^+ R^{-1}(n) / [m_l^+ R^{-1}(n) m_l] \quad l = 1, 2, \dots, L \quad (4.2)$$

来计算 $W_l(n)$ 。

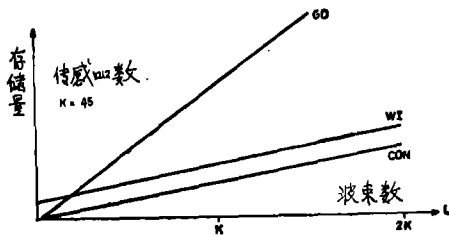


图 4.5 存储量作为传感器数的函数

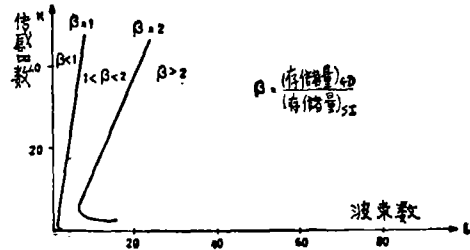


图 4.6 存储量的比(以传感器数和波束数为自变量)

若一个算法能找到合适的参数 μ_{GD} 和 μ_{WI} ，使其平均学习曲线不论在平稳还是在非平稳情况下都比另一个算法的学习曲线低，就说该算法具有的性能比另一个算法好。一种简单的模拟方法是保持数据平稳，然后由初始条件开始自适应，经瞬态(非平稳状况)过渡到稳态(平稳状况)。普通波束形成器的权重矢量可以用于初始条件。

用有 10 个传感器的线阵来作这个模拟，传感器的间隔以比率 1.1 几何递增。最短的间隔和波长的比是 0.3。在 30° 方向和 40° 方向分别有两个点源组成噪声场。噪声场的横向相关半径是阵长的 3 倍。另外还有空间非相干噪声。在传感器上三个场源的相对强度是 10:3:1。

通过这一特例，发现 Woodbury 算法比梯度算法要好。图 4.7 画出学习曲线，并标出了平稳状态的平均误差和标准偏差。图 4.8 上画出的是迭代 1、10、50 和 200 次后的响应。然而，应该特别注意，不要把这么一个模拟实验的结论随意推广，特别是不要推广到传感器数目大得多的情况中去。

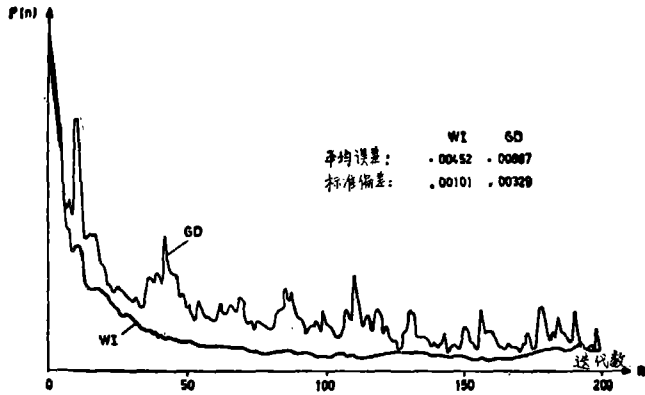


图 4.7 学习曲线

3. 灵活性

从(2.9)式看到，权重矢量 W 依赖于聚焦矢量 m 。因此，若要改变 m ，用 GD 算法进行调整时必须经过瞬态。用 WI 算法就没有这样的情况，这是因为 R^{-1} 与 m 无关，同时这个算法中自适应调整的另一量 $b = m^+ R^{-1} m$ 必须随时直接计算出来。

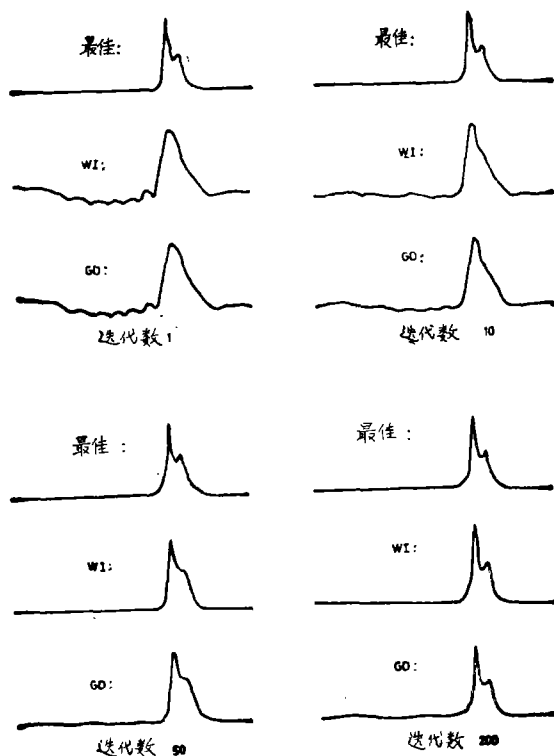


图 4.8 响应的发展过程

五、结 论

和梯度下降算法比较起来, 不论从对计算机的要求、性能还是灵活性来看, Woodbury 恒等算法都有一些很有趣的性质。因此, 对它应该做比公开文献上所显示的情况更为深入的理论研究。不重视这个算法是不应该的。

参 考 文 献

- [1] H. Cox, "Sensitivity Considerations in Adaptive Beamforming", in: Signal Processing, Proc. NATO Advanced Study Institute, 1973.
- [2] N. L. Owsley, "A Recent Trend in Adaptive Spatial Processing for Sensor Arrays: Constrained Adaptation", in: Signal Processing, See ref. 1.
- [3] E. B. Lunde, "Comparison of Two Methods for Constrained Adaptive Beamforming Employing an Auxiliary Array", NUSC Technical Memorandum No. TC-6-75, 1975
- [4] A. Sagehaug, "Analysis of Some Numerical Problems in Adaptive Beamforming," NDRE Tech. Memo., 1975.

(卢 传译 宋健宁校)