

以上を連立方程式として、 $r(0), r(1), r(2)$ を求める。具体的には次のような結果を得る。

$$r(0) = \left(\frac{1+a_2}{1-a_2} \right) \frac{\sigma_v^2}{(1+a_2)^2 - a_1^2} \tag{3.118}$$

$$r(1) = \frac{-a_1}{1+a_2} r(0) \tag{3.119}$$

$$r(2) = \left(-a_2 + \frac{a_1^2}{1+a_2} \right) r(0) \tag{3.120}$$

第4章 最急降下法による適応FIRフィルタ

4.1 適応FIRフィルタの構成

4.1.1 回路構成と方程式

本書では図4.1に示すFIR形の適応フィルタを対象とする。適応フィルタとしてはフィードバックを含む巡回形適応フィルタもあるが安定性の問題もあり、実用的にはFIR形適応フィルタが主に用いられている。

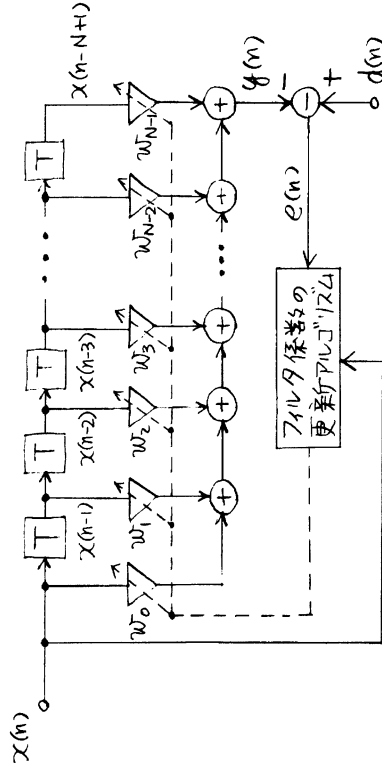


図 4.1: 適応FIRフィルタの構成

入力信号が $x(n)$ でフィルタ係数 w_i への入力が $x(n-i)$ となる。出力 $y(n)$ に対して所望応答 $d(n)$ が与えられ、これらの差が出力誤差 $e(n)$ となる。入力信号 $x(n-i)$ と誤差 $e(n)$ を用いてフィルタ係数 w_i が更新される。

入力信号, 出力信号及び誤差関数は以下のように表される.

$$y(n) = \sum_{i=0}^{N-1} w_i x(n-i) \quad \text{線形和 (結合)} \quad (4.1)$$

$$e(n) = d(n) - y(n) = d(n) - \mathbf{w}^H(n) \mathbf{x}(n) \quad (4.2)$$

$$\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n+N-1)]^T \quad (4.3)$$

$$\mathbf{w}(n) = [w_0(n), w_1(n), \dots, w_{N-1}(n)]^T \quad (4.4)$$

$$\begin{aligned} J(n) &= E[e(n)e^*(n)] \quad \text{二乗平均誤差: Mean Squared Error} \\ &= \sigma_d^2 - \mathbf{w}^H(n) \mathbf{p} - \mathbf{p}^H \mathbf{w}(n) + \mathbf{w}^H(n) \mathbf{R} \mathbf{w}(n) \end{aligned} \quad (4.5)$$

4.1.2 ウィナーフィルタと適応フィルタ

ウィナーフィルタ

$J(n)$ は $\mathbf{w}(n)$ の2次形式であるから, 多次元放物曲面となり, 唯一の極小値を持ちこれが最小値となる. 従って, $\mathbf{w}(n)$ の最適解は

$$\frac{dJ(n)}{d\mathbf{w}(n)} = 0 \quad (4.6)$$

により得られる. この式は次の正規方程式となる.

$$\mathbf{R} \mathbf{w} = \mathbf{p} \quad (4.7)$$

上式をみたす \mathbf{w} を \mathbf{w}_0 とすると, $J(n)$ の最小値は次のようになる.

$$J_{\min} = \sigma_d^2 - \mathbf{p}^H \mathbf{w}_0 \quad (4.8)$$

適応フィルタとは

\mathbf{w} の適当な初期値からスタートして, $J(n)$ の誤差特性曲面 (error performance surface) 上を J_{\min} に向かって逐次修正し, 最終的に $J(n) = J_{\min}$, すなわち $\mathbf{w}(n) = \mathbf{w}_0$ に収束させるものである. このとき, 適応フィルタの出力は目標値に等しくなる, $y(n) = d(n)$. $d(n)$ の決め方により, システム推定, 予測, 等化等を行うことができる.

4.2. 最急降下法

4.2 最急降下法

4.2.1 最急降下法の原理

フィルタ係数 $\mathbf{w}(n)$ を最適値 \mathbf{w}_0 に向かって修正する学習アルゴリズムとして最急降下法 (method of steepest descent) がある. このアルゴリズムでは, $J(n)$ が最小点 (谷底) に向かって斜度を直線的に滑り降りる (直滑降). 最急降下法によるフィルタ係数修正の様子を図4.2に示す. $\mathbf{w}(0)$ が初期値であり, この値に對する $J(n)$ の勾配が $\nabla J(n) = b/a$ である. これに $\mu/2$ をかけた値だけ減少して次の $\mathbf{w}(1)$ が得られる. \mathbf{w}_0 に近づくに連れて勾配が小さくなり, 修正量も減少する. $\mathbf{w}(n) = \mathbf{w}_0$ では勾配=0となり, 修正量も零となり学習が止まる.

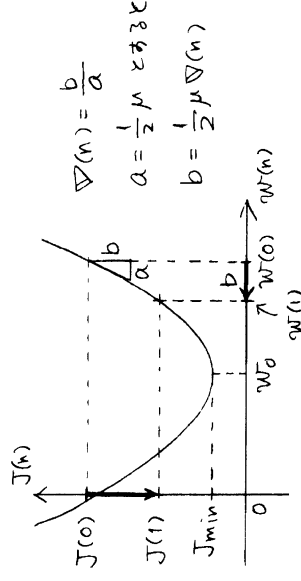


図 4.2: 最急降下法によるフィルタ係数の更新

4.2.2 最急降下法の更新手順

更新手順を以下に示す.

1. 初期値設定: フィルタ係数の初期値を設定する.

$$\text{例えば } \mathbf{w}(0) = \mathbf{0} \quad (4.9)$$

2. 勾配計算: 時刻 n のフィルタ係数 $\mathbf{w}(n)$ を用いて勾配ベクトル (gradient vector) $dJ(n)/d\mathbf{w}(n)$ を計算する.

3. フィルタ係数更新： 時刻 $n+1$ におけるフィルタ係数を求める (更新) .

$$\text{基本式} \quad \mathbf{w}(n+1) = \mathbf{w}(n) - \frac{1}{2} \mu \frac{dJ(n)}{d\mathbf{w}(n)}, \quad \mu > 0 \quad (4.10)$$

4. 誤差の評価：

$$\hat{J}(n) = \frac{1}{M} \sum_{i=0}^{M-1} |e(n-i)|^2 \quad (4.11)$$

μ はフィルタ係数の修正量の大きさを決めるパラメータであり, ステップサイズ (Step size) や学習係数 (Learning rate) と呼ばれる. μ がある条件を満足すれば $\mathbf{w}(n)$ は振動することなく \mathbf{w}_o に一様に収束する.

4.2.3 勾配の統計的な表現

式 (4.5) の $J(n)$ を $\mathbf{w}(n)$ で微分することにより勾配の統計的な表現を得る.

$$\frac{dJ(n)}{d\mathbf{w}(n)} = -2\mathbf{p} + 2\mathbf{R}\mathbf{w}(n) \quad (4.12)$$

これを式 (4.10) に代入する.

$$\begin{aligned} \mathbf{w}(n+1) &= \mathbf{w}(n) + \mu(\mathbf{p} - \mathbf{R}\mathbf{w}(n)) \\ &= (\mathbf{I} - \mu\mathbf{R})\mathbf{w}(n) + \mu\mathbf{p} \end{aligned} \quad (4.13)$$

\mathbf{R}, \mathbf{p} が既知であるとする, 上式よりフィルタ係数が更新できる. しかし, \mathbf{R}, \mathbf{p} は集合平均という統計量であり予め知る事ができない場合が多い. また, 非定常過程では \mathbf{R}, \mathbf{p} は一定ではない. このため, 実際の適応フィルタでは, (1) 近似的な勾配ベクトルを用いる.

(2) \mathbf{R}, \mathbf{p} を逐次的に計算する.

などの方法を用いる. 詳しくは第5章以降で述べる.

4.3 最急降下法の収束性

4.3.1 フィルタ係数の収束性解析

$\mathbf{w}(n)$ が \mathbf{w}_o に収束するための条件を求める.

$$\lim_{n \rightarrow \infty} \mathbf{w}(n) \rightarrow \mathbf{w}_o \quad (4.14)$$

式 (4.13) の関係は図 4.3 のように表される. この回路の安定性はループ利得 $\mathbf{I} - \mu\mathbf{R}$

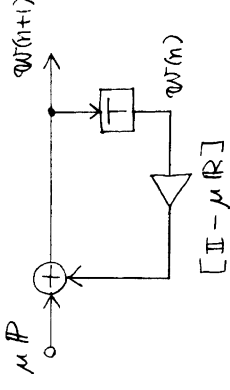


図 4.3: フィルタ係数の更新—ブロック図—

で決まる. \mathbf{R} が既に与えられているとすると, μ に対して安定性条件が課される. $\mathbf{w}(n)$ と \mathbf{w}_o の誤差を

$$\mathbf{c}(n) = \mathbf{w}(n) - \mathbf{w}_o \quad (4.15)$$

とし, $\mathbf{c}(n)$ の時間的な変化を調べることにより, μ に対する条件を求める. まず, 式 (4.13) を変形する. 両辺に $-\mathbf{w}_o$ を加える.

$$\begin{aligned} \mathbf{w}(n+1) - \mathbf{w}_o &= (\mathbf{I} - \mu\mathbf{R})\mathbf{w}(n) - \mathbf{w}_o + \mu\mathbf{p} \\ &= (\mathbf{I} - \mu\mathbf{R})(\mathbf{w}(n) - \mathbf{w}_o) + \mu(\mathbf{p} - \mathbf{R}\mathbf{w}_o) \end{aligned} \quad (4.16)$$

$\mathbf{R}\mathbf{w}_o = \mathbf{p}$ を考慮し, 式 (4.15) の関係を用いると,

$$\mathbf{c}(n+1) = (\mathbf{I} - \mu\mathbf{R})\mathbf{c}(n) \quad (4.17)$$

を得る. さらに, 第3章で述べた次の関係を考慮する.

$$\mathbf{R} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H$$

$$\mathbf{c}(n+1) = (\mathbf{I} - \mu\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H)\mathbf{c}(n) \quad (4.18)$$

上式の両辺に左側から \mathbf{Q}^H をかける。

$$\begin{aligned} \mathbf{Q}^H\mathbf{c}(n+1) &= \mathbf{Q}^H(\mathbf{I} - \mu\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H)\mathbf{c}(n) \\ &= (\mathbf{Q}^H - \mu\mathbf{Q}^H\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H)\mathbf{c}(n) \\ &= (\mathbf{Q}^H - \mu\mathbf{\Lambda}\mathbf{Q}^H)\mathbf{c}(n) \\ &= (\mathbf{I} - \mu\mathbf{\Lambda})\mathbf{Q}^H\mathbf{c}(n) \end{aligned} \quad (4.19)$$

上式では、 \mathbf{Q} がユニタリ行列であり、 $\mathbf{Q}^H = \mathbf{Q}^{-1}$ が成り立つことを利用している。ここで、新たに

$$\mathbf{v}(n) = \mathbf{Q}^H\mathbf{c}(n) = \mathbf{Q}^H[\mathbf{w}(n) - \mathbf{w}_o] \quad (4.20)$$

を定義する。 $\mathbf{v}(n)$ は $\mathbf{w}(n)$ を \mathbf{w}_o だけ移動し、 \mathbf{Q}^H により回転を加えたものである。このようなベクトル変換をアフィン変換 (Affine projection) と呼ぶ。式 (4.19) より、

$$\mathbf{v}(n+1) = (\mathbf{I} - \mu\mathbf{\Lambda})\mathbf{v}(n) \quad (4.21)$$

を得る。 $\mathbf{\Lambda}$ が対角行列であるから、上式に対する各要素の式は次のようになる。

$$v_k(n+1) = (1 - \mu\lambda_k)v_k(n), k = 0, 1, \dots, N-1 \quad (4.22)$$

これより、初期値 $v_k(0)$ からスタートした $v_k(n)$ の一般式は次のようになる。

$$v_k(n) = (1 - \mu\lambda_k)^n v_k(0) \quad (4.23)$$

$\mathbf{w}(n) \rightarrow \mathbf{w}_o$ となるためには、 $\mathbf{v}(n) \rightarrow 0$ とならなければならない。式 (4.23) において

$$\lim_{n \rightarrow \infty} v_k(n) \rightarrow 0, \quad k = 0, 1, \dots, N-1 \quad (4.24)$$

となるためには、

$$|1 - \mu\lambda_k| < 1, \quad k = 0, 1, \dots, N-1 \quad (4.25)$$

が必要十分条件である。上式はさらに次に次のように分解できる。

$$1 - \mu\lambda_k \geq 0, \quad 1 - \mu\lambda_k < 1 \quad (4.26)$$

$$1 - \mu\lambda_k < 0, \quad -1 + \mu\lambda_k < 1 \quad (4.27)$$

これらを満たし、かつ $0 < \mu$ であることを考慮すると

$$0 < \mu < \frac{2}{\lambda_k} \quad (4.28)$$

さらに、 $1/\lambda_k$ の最小値を考慮すると、次の条件を得る。

$$0 < \mu < \frac{2}{\lambda_{max}} \quad (4.29)$$

これが $\mathbf{w}(n)$ が最適解 \mathbf{w}_o に一樣収束するための必要十分条件である。

$\mathbf{w}(n)$ に着目した収束性解析

式 (4.20) を $\mathbf{Q}^{-1} = \mathbf{Q}^H$ を考慮して $\mathbf{w}(n)$ について解く。

$$\begin{aligned} \mathbf{w}(n) &= \mathbf{w}_o + \mathbf{Q}\mathbf{v}(n) \\ &= \mathbf{w}_o + \sum_{k=0}^{N-1} \mathbf{q}_k v_k(n) \end{aligned} \quad (4.30)$$

上式に式 (4.23) を代入する。要素単位の式は次のようになる。

$$\begin{aligned} w_i(n) &= w_{oi} + \sum_{k=0}^{N-1} q_{ki} v_k(0) (1 - \mu\lambda_k)^n \\ i &= 0, 1, \dots, N-1 \end{aligned} \quad (4.31)$$

従って、 $\mathbf{w}(n)$ が \mathbf{w}_o に一樣収束するための条件は式 (4.29) と同じである。

固有値がりと収束速度

式 (4.23) より、 $v_k(n)$ が零に近づく速度、すなわち収束速度は $|1 - \mu\lambda_k|$ で決まる。固有値 λ_k が小さいほどこの値は大きい。すなわち、収束が遅い。一方、3章で述べたように、固有値の総和は入力信号の平均電力に比例する。従って、入力信号の平均電力が一定であるという条件では、固有値がりがり $\chi = \lambda_{max}/\lambda_{min}$ が大きいほど最小固有値 λ_{min} は小さくなる。すなわち、収束が遅くなる。この様子を図 4.4 に示す。3章でも述べたように、入力信号の相関性が強いほど、すなわち、スペクトルの変化 (最大値/最小値) が大きいほど固有値がりがりが大きくなる。

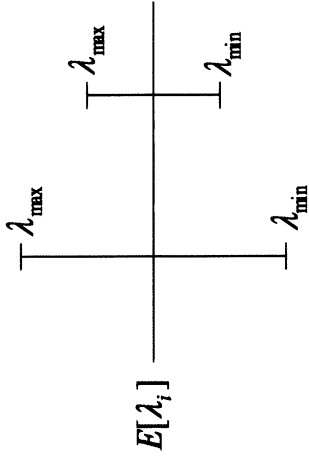


図 4.4: 固有値広がり と 最小固有値の関係.

4.3.2 二乗平均誤差 $J(n)$ の収束性解析

この節では,

- (1) 二乗平均誤差 $J(n)$ の時間的変化 (学習曲線: Learning curve)
 - (2) $J(n)$ が J_{min} に収束するための条件について調べる.
- 式(4.5)を次のように変形する.

$$J(n) = \sigma_d^2 - \mathbf{p}^H \mathbf{w}_o - \mathbf{w}^H(n) \mathbf{p} + \mathbf{p}^H \mathbf{w}_o - \mathbf{p}^H \mathbf{w}(n) + \mathbf{w}^H(n) \mathbf{R} \mathbf{w}(n) \quad (4.32)$$

さらに,

$$\begin{aligned} J_{min} &= \sigma_d^2 - \mathbf{p}^H \mathbf{w}_o \\ \mathbf{R} \mathbf{w}_o &= \mathbf{p} \\ \mathbf{R}^H &= \mathbf{R} \end{aligned}$$

を考慮すると, 次式を得る.

$$J(n) = J_{min} - \mathbf{w}^H(n) \mathbf{R} \mathbf{w}_o + (\mathbf{R} \mathbf{w}_o)^H \mathbf{w}_o - (\mathbf{R} \mathbf{w}_o)^H \mathbf{w}(n) + \mathbf{w}^H(n) \mathbf{R} \mathbf{w}(n) \quad (4.33)$$

$$= J_{min} - \mathbf{w}^H(n) \mathbf{R} \mathbf{w}_o + \mathbf{w}_o^H \mathbf{R} \mathbf{w}_o - \mathbf{w}_o^H \mathbf{R} \mathbf{w}(n) + \mathbf{w}^H(n) \mathbf{R} \mathbf{w}(n) \quad (4.34)$$

$$= J_{min} + (\mathbf{w}(n) - \mathbf{w}_o)^H \mathbf{R} (\mathbf{w}(n) - \mathbf{w}_o) \quad (4.35)$$

さらに, 式(4.20)と $\mathbf{R} = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^H$ を考慮すると

$$\begin{aligned} J(n) &= J_{min} + \mathbf{v}^H(n) \mathbf{\Lambda} \mathbf{v}(n) \\ \mathbf{v}(n) &= \mathbf{Q}^H (\mathbf{w}(n) - \mathbf{w}_o) \end{aligned} \quad (4.36)$$

となる. この式を誤差特性曲面の**最小形式**(canonical form)という. $\mathbf{\Lambda}$ が対角行列であるから, 次のように分解される.

$$\begin{aligned} J(n) &= J_{min} + \sum_{k=1}^N \lambda_k v_k(n) v_k^*(n) \\ &= J_{min} + \sum_{k=1}^N \lambda_k |v_k(n)|^2 \end{aligned} \quad (4.37)$$

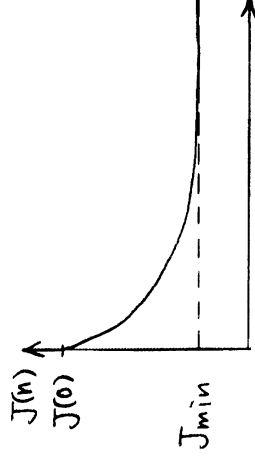
$v_k(n)$ に式(4.23)を代入すると, 時間 n における誤差は次のように表される.

$$J(n) = J_{min} + \sum_{k=1}^N \lambda_k (1 - \mu \lambda_k)^{2n} |v_k(0)|^2 \quad (4.38)$$

従って, この場合も μ が式(4.29)の条件を満たせば, $J(n)$ は J_{min} に一様収束する.

$$\lim_{n \rightarrow \infty} J(n) \rightarrow J_{min} \quad (4.39)$$

この様子を図 4.5 に示す.

図 4.5: 学習曲線: 一様収束する $J(n)$.

統計量の具体的な計算法

誤差関数を $J(n) = E[e(n)e^*(n)]$ としたとき, 正確な勾配を用いれば学習は一樣

収束する。しかし、正確な勾配は統計量であり、実際に求めることはできない。そこで、実際には

1. 勾配法では、 $J(n) = E[e(n)e^*(n)]$ を瞬時値 $\hat{J}(n) = e(n)e^*(n)$ で置き換える。…LMS, NLMS アルゴリズム

2. 最小2乗法では、 $\mathbf{R} = E[\mathbf{x}(n)\mathbf{x}^H(n)]$ の代わりに

$$\Phi(n) = \sum_{i=1}^n \lambda^{n-i} \mathbf{x}(n)\mathbf{x}^H(n) \quad (4.40)$$

を用いる。…LS, RLS アルゴリズム

第5章 最小二乗平均アルゴリズム

5.1 最小二乗平均アルゴリズムの導出

5.1.1 統計量としての勾配

4章で述べたように、勾配ベクトルが

$$\frac{dJ(n)}{d\mathbf{w}(n)} = -2\mathbf{p} + 2\mathbf{R}\mathbf{w}(n) \quad (5.1)$$

から正確に推定され、ステップサイズ μ が式 (4.29) の収束範囲に選ばれたならば、式 (4.10) の最急降下法によりフィルタ係数 $\mathbf{w}(n)$ は \mathbf{w}_0 に、二乗平均誤差 $J(n)$ は J_{min} にそれぞれ一様収束する。

しかし、 \mathbf{R}, \mathbf{p} は次に示すように期待値であり、実際には予め知る (測定する) ことが難しい。また、非定常過程ではこれらは時間とともに変化するので更新する必要がある。従って、実際にはこれらの近似値が用いられる。

$$\mathbf{R} = E[\mathbf{x}(n)\mathbf{x}^H(n)] \quad (5.2)$$

$$\mathbf{p} = E[\mathbf{x}(n)d^*(n)] \quad (5.3)$$

5.1.2 統計的勾配の近似

もっとも簡単な近似は期待値の代わりに瞬時値を用いることである。このような学習アルゴリズムを**最小二乗平均** (Least Mean Square: LMS) アルゴリズムという。 $J(n), \mathbf{R}, \mathbf{p}$ に対する瞬時値による近似を $\hat{J}(n), \hat{\mathbf{R}}(n), \hat{\mathbf{p}}(n)$ とする。

$$\hat{J}(n) = e(n)e^*(n) \quad (5.4)$$

$$\hat{\mathbf{R}}(n) = \mathbf{x}(n)\mathbf{x}^H(n) \quad (5.5)$$

$$\hat{\mathbf{p}}(n) = \mathbf{x}(n)d^*(n) \quad (5.6)$$

勾配に対する近似は式(5.1)より, 次のようになる.

$$\begin{aligned} \frac{d\hat{J}(n)}{d\hat{\mathbf{w}}(n)} &= -2\hat{\mathbf{p}}(n) + 2\hat{\mathbf{R}}\hat{\mathbf{w}}(n) \\ &= -2\mathbf{x}(n)d^*(n) + 2\mathbf{x}(n)\mathbf{x}^H(n)\hat{\mathbf{w}}(n) \\ &= -2\mathbf{x}(n)[d^*(n) - \mathbf{x}^H(n)\hat{\mathbf{w}}(n)] \end{aligned} \quad (5.7)$$

$$(5.8)$$

$\hat{\mathbf{w}}(n)$ は近似勾配を用いて更新されるフィルタ係数である.

5.1.3 フィルタ係数の更新式

$\hat{\mathbf{w}}(n)$ の更新式は最急降下法の式

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) - \frac{1}{2}\mu \frac{d\hat{J}(n)}{d\hat{\mathbf{w}}(n)} \quad (5.9)$$

に式(5.8)を代入する.

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu\mathbf{x}(n)[d^*(n) - \mathbf{x}^H(n)\hat{\mathbf{w}}(n)] \quad (5.10)$$

フィルタ係数として $\hat{\mathbf{w}}(n)$ を用いたときのフィルタ出力を $\hat{y}(n)$, 出力誤差を $\hat{e}(n)$ とする.

$$\hat{y}(n) = \hat{\mathbf{w}}^H(n)\mathbf{x}(n) \quad (5.11)$$

$$\hat{e}(n) = d(n) - \hat{y}(n) \quad (5.12)$$

$$= d(n) - \hat{\mathbf{w}}^H(n)\mathbf{x}(n) \quad (5.13)$$

以上より, $\hat{\mathbf{w}}(n)$ の更新式(5.10)は次のようになり, 修正量は入力信号と出力誤差で決まる.

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu\mathbf{x}(n)\hat{e}^*(n) \quad (5.14)$$

$$\hat{w}_k(n+1) = \hat{w}_k(n) + \mu x(n-k)\hat{e}^*(n), \quad k = 0, 1, \dots, N-1 \quad (5.15)$$

$x(n-k)$ は $w_k(n)$ に対する入力信号である. 図5.1にFIR形適応フィルタとフィルタ係数更新のブロック図を示す.

計算量

LMSアルゴリズムでは, N 個のフィルタ係数を1回更新するのに必要な計算は $2N$ 回の乗算と N 回の加算である.

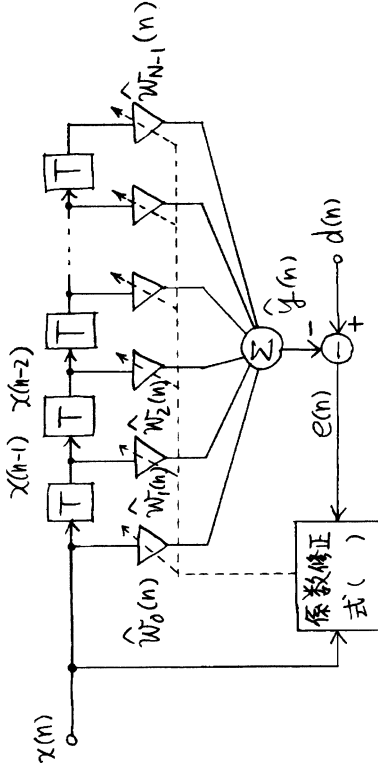


図 5.1: FIR 形適応フィルタと係数更新のブロック図

5.2 フィルタ係数の収束性

期待値を瞬時値で近似したLMSアルゴリズムにおけるフィルタ係数の収束性について調べる.

5.2.1 $E[\hat{\mathbf{w}}(n)]$ の収束性

式(5.10)の両辺から \mathbf{w}_o を引き, さらに

$$\hat{\mathbf{e}}(n) = \hat{\mathbf{w}}(n) - \mathbf{w}_o \quad (5.16)$$

と変形する.

$$\hat{\mathbf{w}}(n+1) - \mathbf{w}_o = \hat{\mathbf{w}}(n) - \mathbf{w}_o + \mu\mathbf{x}(n)[d^*(n) - \mathbf{x}^H(n)\hat{\mathbf{w}}(n)] \quad (5.17)$$

$$\begin{aligned} \hat{\mathbf{e}}(n+1) &= [\mathbf{I} - \mu\mathbf{x}(n)\mathbf{x}^H(n)]\hat{\mathbf{e}}(n) \\ &+ \mu[\mathbf{x}(n)d^*(n) - \mathbf{x}(n)\mathbf{x}^H(n)\mathbf{w}_o] \end{aligned} \quad (5.18)$$

ここで、両辺の期待値をとる。

$$\begin{aligned}
 E[\hat{\mathbf{c}}(n+1)] &= E[(\mathbf{I} - \mu \mathbf{x}(n) \mathbf{x}^H(n)) \hat{\mathbf{c}}(n)] \\
 &+ \mu E[\mathbf{x}(n) \mathbf{d}^*(n) - \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{w}_o] \\
 &= (\mathbf{I} - \mu E[\mathbf{x}(n) \mathbf{x}^H(n)]) E[\hat{\mathbf{c}}(n)] \\
 &+ \mu \{E[\mathbf{x}(n) \mathbf{d}^*(n)] - E[\mathbf{x}(n) \mathbf{x}^H(n)] \mathbf{w}_o\} \\
 &= (\mathbf{I} - \mu \mathbf{R}) E[\hat{\mathbf{c}}(n)] + \mu (\mathbf{p} - \mathbf{R} \mathbf{w}_o)
 \end{aligned} \tag{5.19}$$

式 (5.19) では、

$$E[\mathbf{x}(n) \mathbf{x}^H(n) \hat{\mathbf{c}}(n)] = E[\mathbf{x}(n) \mathbf{x}^H(n)] E[\hat{\mathbf{c}}(n)] \tag{5.21}$$

としている。\$\mathbf{x}(n)\$ は音声などの信号であるのに対して、\$\hat{\mathbf{c}}(n)\$, すなわち \$\hat{\mathbf{w}}(n)\$ は未知システムのインパルス応答であり、これらは独立である場合が多い。さらに、\$\mathbf{R} \mathbf{w}_o = \mathbf{p}\$ であるから、式 (5.20) は

$$E[\hat{\mathbf{c}}(n+1)] = (\mathbf{I} - \mu \mathbf{R}) E[\hat{\mathbf{c}}(n)] \tag{5.22}$$

となる。この関係は式 (4.17) において \$\mathbf{c}(n)\$ を \$E[\hat{\mathbf{c}}(n)]\$ で置き換えた形になっている。従って、LMS アルゴリズムにおける収束性の議論は最急降下法の結果を \$E[\hat{\mathbf{w}}(n)]\$ に適用することにより求められる。

$$\lim_{n \rightarrow \infty} E[\hat{\mathbf{c}}(n)] \rightarrow 0 \tag{5.23}$$

$$0 < \mu < \frac{2}{\lambda_{max}}$$

以上のように、LMS アルゴリズムではステップ \$\mu\$ が式 (4.29) の条件を満たすとき \$E[\hat{\mathbf{w}}(n)]\$ が \$\mathbf{w}_o\$ に一樣収束する。

5.2.2 \$\hat{\mathbf{w}}(n)\$ の収束性

\$\hat{\mathbf{c}}(n)\$ の相関行列を考える。

$$\hat{\mathbf{C}}(n) = E[\hat{\mathbf{c}}(n) \hat{\mathbf{c}}^H(n)] \tag{5.24}$$

5.3 二乗平均誤差の収束性

式 (5.18) を上式に代入することにより次式を得る。

$$\begin{aligned}
 \hat{\mathbf{C}}(n+1) &= \hat{\mathbf{C}}(n) - \mu [\mathbf{R} \hat{\mathbf{C}}(n) + \hat{\mathbf{C}}(n) \mathbf{R}] \\
 &+ \mu^2 \mathbf{R} \text{tr}[\mathbf{R} \hat{\mathbf{C}}(n)] + \mu^2 J_{min} \mathbf{R}
 \end{aligned} \tag{5.25}$$

上式では、\$\mu^2 J_{min} \mathbf{R} \neq 0\$ であるため \$\hat{\mathbf{C}}(n+1) = \hat{\mathbf{C}}(n) = 0\$ という解はない。すなわち、\$\hat{\mathbf{C}}(n)\$ は零に収束しない。\$\hat{\mathbf{C}}(n)\$ を小さくするためには、\$\mu\$ を小さくする必要がある。

以上の結果より、\$\hat{\mathbf{w}}(n)\$ は \$\mathbf{w}_o\$ には収束せず、その周りで振動する。\$\hat{\mathbf{w}}(n)\$ を \$\mathbf{w}_o\$ に近づけるためには \$\mu\$ を十分に小さくする必要がある。

5.3 二乗平均誤差の収束性

5.3.1 \$J(n)\$ の収束性

式 (4.35) より

$$\begin{aligned}
 J(n) &= J_{min} + (\hat{\mathbf{w}}(n) - \mathbf{w}_o)^H \mathbf{R} (\hat{\mathbf{w}}(n) - \mathbf{w}_o) \\
 &= J_{min} + \hat{\mathbf{c}}(n)^H \mathbf{R} \hat{\mathbf{c}}(n)
 \end{aligned} \tag{5.26}$$

前節で述べたように \$\hat{\mathbf{c}}(n)\$ は零に収束しないから

$$\hat{\mathbf{c}}(n)^H \mathbf{R} \hat{\mathbf{c}}(n) > 0 \tag{5.27}$$

となる。すなわち、\$J(n)\$ は \$J_{min}\$ に収束できない。超過分を \$J_{ex}(n)\$ とする。

$$J_{ex}(n) = \hat{\mathbf{c}}(n)^H \mathbf{R} \hat{\mathbf{c}}(n) \tag{5.28}$$

$$J(n) = J_{min} + J_{ex}(n) \tag{5.29}$$

5.3.2 \$E[J(n)]\$ の収束性

式 (5.29) の両辺の期待値をとる

$$E[J(n)] = J_{min} + E[J_{ex}(n)] \tag{5.30}$$

\$E[J_{ex}(n)]\$ の収束性について考える。

$$E[J_{ex}(n)] = E[\hat{\mathbf{c}}(n)^H \mathbf{R} \hat{\mathbf{c}}(n)] = \text{tr}[\mathbf{R} \hat{\mathbf{C}}(n)] \tag{5.31}$$

一方, 式 (3.71) より

$$\mathbf{R} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H \quad (5.32)$$

$\mathbf{\Lambda}$ は固有値を対角要素とする対角行列, \mathbf{Q} は固有ベクトルを列ベクトルとするユニタリ行列である. さらに,

$$\mathbf{\Gamma}(n) = \mathbf{Q}^H \hat{\mathbf{C}}(n) \mathbf{Q} \quad (5.33)$$

より求める $\mathbf{\Gamma}$ を用いて

$$\hat{\mathbf{C}}(n) = \mathbf{Q}\mathbf{\Gamma}(n)\mathbf{Q}^H \quad (5.34)$$

とおく. $\mathbf{\Gamma}$ は一般に対角行列ではない. 式 (5.31) に式 (5.32), (5.34) を代入する.

$$\begin{aligned} E[J_{ex}(n)] &= \text{tr}[\mathbf{R}\hat{\mathbf{C}}(n)] \\ &= \text{tr}[\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H \mathbf{Q}\mathbf{\Gamma}(n)\mathbf{Q}^H] \\ &= \text{tr}[\mathbf{Q}\mathbf{\Lambda}\mathbf{\Gamma}(n)\mathbf{Q}^H] \\ &= \text{tr}[\mathbf{Q}^H \mathbf{Q}\mathbf{\Lambda}\mathbf{\Gamma}(n)] \\ &= \text{tr}[\mathbf{\Lambda}\mathbf{\Gamma}(n)] \end{aligned} \quad (5.35) \quad (5.36)$$

上式では, $\mathbf{Q}^H \mathbf{Q} = \mathbf{I}$, $\text{tr}[\mathbf{A}\mathbf{B}] = \text{tr}[\mathbf{B}\mathbf{A}]$ であることを利用している. さらに, 行列の要素を用いて

$$E[J_{ex}(n)] = \sum_{i=1}^N \lambda_i \gamma_i(n) \quad (5.37)$$

と表される. γ_i は $\mathbf{\Gamma}$ の対角要素である.

$\gamma(n)$ の収束性

式 (5.25) の $\hat{\mathbf{C}}(n)$ に対する漸化式に左側から \mathbf{Q}^H を, 右側から \mathbf{Q} をかけて, $\mathbf{\Gamma}$ と $\mathbf{\Lambda}$ を用いた式に変形する.

$$\begin{aligned} \mathbf{\Gamma}(n+1) &= \mathbf{\Gamma}(n) - \mu[\mathbf{\Lambda}\mathbf{\Gamma}(n) + \mathbf{\Gamma}(n)\mathbf{\Lambda}] \\ &\quad + \mu^2 \mathbf{\Lambda} \text{tr}[\mathbf{\Lambda}\mathbf{\Gamma}(n)] + \mu^2 J_{min} \mathbf{\Lambda} \end{aligned} \quad (5.38)$$

$\mathbf{\Gamma}(n)$ の対角要素 $\gamma_i(n)$ について漸化式を求める.

$$\begin{aligned} \gamma_i(n+1) &= \gamma_i(n) - 2\mu\lambda_i\gamma_i(n) \\ &\quad + \mu^2 \lambda_i \sum_{j=1}^N \lambda_j \gamma_j(n) + \mu^2 J_{min} \lambda_i \\ i &= 1, 2, \dots, N \end{aligned} \quad (5.39)$$

さらにベクトルと行列を用いて表す.

$$\gamma(n+1) = \mathbf{B}\gamma(n) + \mu^2 J_{min} \boldsymbol{\lambda} \quad (5.40)$$

$$\gamma(n) = [\gamma_1(n), \gamma_2(n), \dots, \gamma_N(n)]^T \quad (5.41)$$

$$\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_N]^T \quad (5.42)$$

$$\mathbf{B} = \{b_{ij}\} \quad N \times N \text{ 対称行列}$$

$$b_{ij} = \begin{cases} (1 - \mu\lambda_i)^2, & i = j \\ \mu^2 \lambda_i \lambda_j, & i \neq j \end{cases} \quad (5.43)$$

\mathbf{B} は対称行列でその要素は正の実数である $b_{ij} = b_{ji} > 0$. 式 (5.40) より, $\gamma(n)$ は初期値 $\gamma(0)$ を仮定すると次の一般式で与えられる.

$$\gamma(n) = \mathbf{B}^n \gamma(0) + \mu^2 J_{min} \sum_{i=0}^{n-1} \mathbf{B}^i \boldsymbol{\lambda} \quad (5.44)$$

上式を次の関係を用いて変形する.

$$(\mathbf{I} - \mathbf{B}^n) = (\mathbf{I} - \mathbf{B})(\mathbf{I} + \mathbf{B} + \mathbf{B}^2 + \dots + \mathbf{B}^{n-1}) \quad (5.45)$$

$$\sum_{i=0}^{n-1} \mathbf{B}^i = \frac{\mathbf{I} - \mathbf{B}^n}{\mathbf{I} - \mathbf{B}} \quad (5.46)$$

$$\begin{aligned} \gamma(n) &= \mathbf{B}^n [\gamma(0) - \mu^2 J_{min} (\mathbf{I} - \mathbf{B})^{-1} \boldsymbol{\lambda}] \\ &\quad + \mu^2 J_{min} (\mathbf{I} - \mathbf{B})^{-1} \boldsymbol{\lambda} \end{aligned} \quad (5.47)$$

上式の第1項は n によって変化する過渡成分で, 第2項が n によって変化しない定常成分である.

$\gamma(n)$ の過渡成分が零に収束する条件

\mathbf{B} の固有値を対角要素とする対角行列を \mathbf{F} , 固有ベクトルを列ベクトルとする

行列を \mathbf{G} とすると

$$\mathbf{B} = \mathbf{G}\mathbf{F}\mathbf{G}^H \quad (5.48)$$

$$\mathbf{B}^n = \mathbf{G}\mathbf{F}^n\mathbf{G}^H \quad (5.49)$$

であり, $\gamma(n)$ の繰り返し回数 n に対する変化は \mathbf{B} の固有値の n 乗で決まる. 従って, $n \rightarrow \infty$ に対して式 (5.47) の過渡成分が零に収束するための必要十分条件は \mathbf{B} の固有値 f_i の絶対値が 1 未満である.

$$|f_i| < 1, \quad i = 1, 2, \dots, N \quad (5.50)$$

このとき, $\gamma(n)$ は第 2 項に収束する.

$$\lim_{n \rightarrow \infty} \gamma(n) \rightarrow \mu^2 J_{\min}(\mathbf{I} - \mathbf{B})^{-1} \boldsymbol{\lambda} \quad (5.51)$$

$$|f_i| < 1, \quad i = 1, 2, \dots, N$$

$|f_i| < 1$ の条件を μ の条件に変換

\mathbf{B} の固有値を f , これに対する固有ベクトルを \mathbf{g} とする.

$$\mathbf{B}\mathbf{g} = f\mathbf{g} \quad (5.52)$$

上式で, $\mathbf{g} = [g_1, g_2, \dots, g_N]^T$ とし, \mathbf{B} に式 (5.43) の b_{ij} を代入する.

$$g_i - 2\mu\lambda_i g_i + \mu^2 \left(\sum_{j=1}^N \lambda_j g_j \right) \lambda_i = f g_i \quad (5.53)$$

$$i = 1, 2, \dots, N$$

上式において, 収束条件の境界である $f = 1$ に対する μ を求める. そのために, $f = 1$ を代入し, 両辺を i について加算する.

$$2\mu \sum_{i=1}^N \lambda_i g_i = \mu^2 \left(\sum_{j=1}^N \lambda_j g_j \right) \sum_{i=1}^N \lambda_i \quad (5.54)$$

これより,

$$\mu = \frac{2}{N} \sum_{i=1}^N \lambda_i \quad (5.55)$$

$0 < \mu$ であり, また相関行列 \mathbf{R} の固有値は非負 ($0 \leq \lambda_i$) であるから, μ に対する安定条件は

$$0 < \mu < \frac{2}{\sum_{i=1}^N \lambda_i} \quad (5.56)$$

となる.

5.3.3 μ と入力電力の関係

固有値の総和は対角要素の和 (トレース) に等しい.

$$\sum_{i=1}^N \lambda_i = N r(0) \quad (5.57)$$

$$r(0) = \sigma_x^2 \quad \text{但し } E[x(n)] = 0 \quad (5.58)$$

であるから, 式 (5.56) は

$$0 < \mu < \frac{2}{N\sigma_x^2} \quad (5.59)$$

となる. σ_x^2 は入力信号 $x(n)$ の平均電力である. すなわち, ステップサイズ μ に対する収束条件は固有値を計算¹する必要が無く, 入力信号の平均電力²で規定できる.

5.3.4 $E[\hat{\mathbf{w}}(n)]$ と $E[J(n)]$ の収束条件

$E[\hat{\mathbf{w}}(n)]$ は $0 < \mu < 2/\lambda_{\max}$ の条件下で \mathbf{w}_0 に一樣収束する. このとき, $\hat{\mathbf{w}}(n)$ は \mathbf{w}_0 の回りで振動する. $J(n)$ も振動する. 一方, $E[J(n)]$ は $0 < \mu < 2/\sum_{i=1}^N \lambda_i$ の条件下で $J_{\min} + E[J_{\text{ex}}(\infty)]$ に一樣収束する. 後者の方が条件が厳しい. 言い換えると, $E[\hat{\mathbf{w}}(n)]$ が一樣収束しても, そのとき $E[J(n)]$ が一樣収束しているとは限らない. 逆に $E[J(n)]$ が一樣収束しているときは, $E[\hat{\mathbf{w}}(n)]$ も必ず一樣収束する. 実際に μ は, $E[J_{\text{ex}}(\infty)]$ を小さくすることも考慮して $2/\sum_{i=1}^N \lambda_i$ よりも小さく設定する.

¹ N 次方程式を解く必要がある.

² 定常信号, あるいは短時間区間で定常信号と見なすことができれば時間平均で計算できる.

5.3.5 誤調整

LMS アルゴリズムでは、式 (5.30) に示すようにステップサイズ μ が収束条件を満たし、 $E[J(n)]$ が一様収束しても J_{min} には収束できず、 $E[J_{ex}(\infty)]$ だけの超過分が発生する。 J_{min} と $E[J_{ex}(\infty)]$ の比を**誤調整**(Misadjustment) と呼ぶ。式 (5.40) において $n \rightarrow \infty$ とする。このとき、 μ は式 (5.56) の収束条件を満たすものとする。 $\gamma_i(n)$ はある値に収束するからこれを $\gamma_i(\infty)$ とする。すなわち、式 (5.40) において

$$\gamma_i(n+1) = \gamma_i(n) = \gamma_i(\infty) \quad (5.60)$$

とする。

$$\begin{aligned} \gamma_i(\infty) &= \gamma_i(\infty) - 2\mu\lambda_i\gamma_i(\infty) \\ &+ \mu^2\lambda_i \sum_{j=1}^N \lambda_j\gamma_j(\infty) + \mu^2 J_{min}\lambda_i \end{aligned} \quad (5.61)$$

$\lambda_i\gamma_i(\infty)$ に着目して、 i について加算する。

$$\begin{aligned} 2 \sum_{i=1}^N \lambda_i\gamma_i(\infty) &= \mu \sum_{i=1}^N \lambda_i \sum_{i=1}^N \lambda_i\gamma_i(\infty) \\ &+ \mu J_{min} \sum_{i=1}^N \lambda_i \end{aligned} \quad (5.62)$$

$$\sum_{i=1}^N \lambda_i\gamma_i(\infty) = \frac{\mu J_{min} \sum_{i=1}^N \lambda_i}{2 - \mu \sum_{i=1}^N \lambda_i} \quad (5.63)$$

式 (5.37) より、上式が $E[J_{ex}(\infty)]$ を与える。

$$E[J_{ex}(\infty)] = \frac{\mu J_{min} \sum_{i=1}^N \lambda_i}{2 - \mu \sum_{i=1}^N \lambda_i} \quad (5.64)$$

誤調整は次のようになる。

$$\epsilon = \frac{E[J_{ex}(\infty)]}{J_{min}} = \frac{\mu \sum_{i=1}^N \lambda_i}{2 - \mu \sum_{i=1}^N \lambda_i} \quad (5.65)$$

上式から分かるように、誤調整 ϵ はおよそ μ に比例している。従って、 ϵ を小さくするためには、 μ を収束条件の上限よりも十分に小さくする必要はある。一方、 μ を小さくしすぎると学習が遅くなるという問題もある。実際の応用では、ステップサイズの決め方や制御方法が重要なポイントである。

5.4 LMS アルゴリズムのまとめ

フィルタ次数

フィルタ次数： $N - 1$

フィルタ係数の数 (タップ数)： N

ステップサイズ

$$0 < \mu < \frac{2}{N} = \frac{2}{Nr(0)} = \frac{2}{NE[x(n)x^*(n)]} \quad (5.66)$$

データ

入力信号： $x(n) = [x(n), x(n-1), \dots, x(n-N+1)]^T$

希望応答 (目標値)： $d(n)$

フィルタ係数の更新

初期値： $\hat{w}(0) = \mathbf{0}$

フィルタ出力： $\hat{y}(n) = \hat{w}^H(n)\hat{x}(n)$

誤差： $\hat{e}(n) = d(n) - \hat{y}(n)$

フィルタ係数の更新： $\hat{w}(n+1) = \hat{w}(n) + \mu\hat{x}(n)\hat{e}^*(n)$

誤差評価： $E = \frac{1}{L} \sum_{l=0}^{L-1} |e(n-l)|^2$

L は誤差を評価する区間長。

実際には、目標信号 $d(n)$ 、あるいは、それを出力する未知システムの特性が時間的に変化する場合が多く、フィルタ係数の更新は常時行われる。

5.5 LMS アルゴリズムの変形

5.5.1 正規化 LMS アルゴリズム: Normalized LMS (NLMS)

LMS アルゴリズムでは、フィルタ係数の修正量は $\mu \hat{e}^*(n) \mathbf{x}(n)$ である。 μ が収束条件の範囲内で固定されているとすると、修正量の変動は $\hat{e}^*(n) \mathbf{x}(n)$ に依存する。特に、入力信号の瞬時値によって大きく変動する。この変動を押さえるために μ を小さくすると収束が遅くなるという問題もある。

このように、ステップサイズは収束を保証するためには小さくする必要があるが、小さすぎると収束が遅くなる。そのため、適当な値に設定する必要がある。また、入力信号の時間的な変動にも追従する必要がある。

そこで、修正量における $\mathbf{x}(n)$ をそのノルムの二乗で正規化した $\mathbf{x}(n) / \|\mathbf{x}(n)\|^2$ を用いる。これは見方を変えれば、ステップサイズを μ から $\alpha / \|\mathbf{x}(n)\|^2$ に変更したことに相当する。すなわち、ステップサイズを入力信号の電力に応じて可変することに相当する。

NLMS アルゴリズムは次式で表される。

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \frac{\alpha}{\|\mathbf{x}(n)\|^2} \mathbf{x}(n) \hat{e}^*(n) \quad (5.67)$$

$$\|\mathbf{x}(n)\|^2 = \sum_{k=0}^{N-1} \mathbf{x}(n-k) \mathbf{x}^*(n-k) \quad (5.68)$$

δ は $\|\mathbf{x}(n)\|^2$ が非常に小さくなったときに数値計算を安定化するために用いられる小さい正の定数である。

入力信号が定常信号、または、短時間内で定常信号である場合は

$$\|\mathbf{x}(n)\|^2 = N r(0) = N \sigma_x^2 \quad (5.69)$$

と見なせるので、 α に対する収束条件は式 (5.56), (5.67), (5.69) より

$$0 < \alpha < 2 \quad (5.70)$$

となる。

このように、NLMS アルゴリズムは式 (5.56) の収束条件を満たし、かつ、 σ_x^2 、すなわち入力信号の平均電力の時間的な変動に追従し、常に最適なステップサイズを与えることができる。

5.5.2 クリップ形 LMS アルゴリズム

LMS アルゴリズムの修正量において、 $\mathbf{x}(n)$ の代わりにその符号のみを用いる方法である。符号 (サイン) アルゴリズムとも呼ばれる。適応フィルタとしての性能を大きく劣化させないでフィルタ係数更新における計算を簡単にできる。

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu e(n) \text{sgn}[\mathbf{x}(n)] \quad (5.71)$$

$$\text{sgn}[x] = \begin{cases} +1, & x \geq 0 \\ -1, & x < 0 \end{cases} \quad (5.72)$$

5.5.3 可変ステップサイズ方式

NLMS でも α を可変することができる。ステップサイズを可変する目的は

- 収束を速くする (ステップサイズ \rightarrow 大)。
 - 最終誤差を小さくする (ステップサイズ \rightarrow 小)。
 - 誤差に干渉成分を含む場合はこの影響を小さくする (ステップサイズ \rightarrow 小)。
- 等である。詳細はここでは省略する。

5.6 システム推定のシミュレーション

5.6.1 システム構成

図 5.2 は適応フィルタにより未知システムを推定するブロック図である。

入力信号 $x(n)$ は白色雑音 $v_1(n)$ を 2 次自己回帰 (Auto-regressive: AR) 形回路 $G(z)$ を通して生成される。

$$G(z) = \frac{a_0}{1 + b_1 z^{-1} + b_2 z^{-2}} \quad (5.73)$$

$$b_1 = -2r \cos \theta \quad (5.74)$$

$$b_2 = r^2 \quad (5.75)$$

$$\theta = \pi/2$$

$$r = 0.1, 0.8, 0.9$$

$G(z)$ の極の大きさが r で角度が θ である。シミュレーションでは r を調整して固有値広がりを変える。 r を変えたときの $G(z)$ の振幅特性を図 5.3 に示す。未

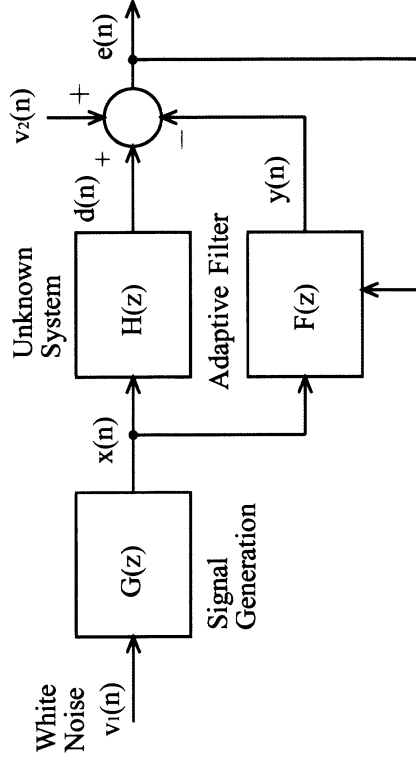


図 5.2: システム推定のシミュレーションを行うブロック図

知システム $H(z)$ は FIR フィルタとし、そのインパルス応答と振幅特性を図 5.4, 5.5 に示す。インパルス応答の長さは 20 サンプルである。適応フィルタ $F(z)$ も FIR フィルタとし、そのインパルス応答の長さも 20 ($N = 20$) サンプルとする。未知システムと適応フィルタの出力の差が誤差であり、この二乗値を最小にするように適応フィルタが学習される。観測雑音として白色の $v_2(n)$ を考慮する。

5.6.2 学習曲線

図 5.6, 5.7 に LMS アルゴリズムによる学習曲線を、図 5.8, 5.9 に NLMS アルゴリズムによる学習曲線を示す。入力信号を変えて 10 回の試行を行い二乗誤差の集合平均をとり、入力信号の平均電力との比で表している (Mean Squared Error in dB)。 k 回目試行の誤差を $e_k(n)$ とする。集合平均は

$$E(n) = \frac{1}{10} \sum_{k=1}^{10} e_k^2(n) \quad (5.76)$$

となる。試行によって白色雑音 $v_1(n)$ のサンプル値は変化するが統計的な性質は変化しない。従って、入力信号 $x(n)$ も統計的な性質は変化しない。また、定常過程であるから時間方向に二乗平均を計算し、これを平均電力とする。これを

σ_x^2 とするとき、二乗平均誤差は次式で与えられる。

$$E_{av}(n) = 10 \log_{10} \frac{E(n)}{\sigma_x^2} \quad [\text{dB}] \quad (5.77)$$

5.6.3 ステップサイズについて

LMS アルゴリズムにおけるステップサイズは次式で表している。

$$\begin{aligned} \mu &= \frac{\mu_0}{N\sigma_x^2} \\ \mu_0 &= 0.7, 0.5 \end{aligned} \quad (5.78)$$

σ_x^2 は入力信号の全サンプルに対する二乗平均である。 $\mu_0 > 0.7$ では、 $\tau = 0.9$ のときに学習が発散した。

NLMS アルゴリズムの α は式 (5.67) で用いられているものである。シミュレーションでは

$$\alpha = 1, 0.5$$

を用いた。いずれの場合も、 τ が大きくなるほど、すなわち固有値広がりが大きくなるほど収束が遅くなることが分かる。

LMS アルゴリズムに比べて NLMS アルゴリズムの方が安定している。これは、LMS アルゴリズムではステップサイズが固定されているのに対して、NLMS アルゴリズムでは現時点から適応フィルタ長である N サンプル遡った区間で $\|\mathbf{x}(n)\|^2$ を計算しているので入力信号のレベル変動に追従できるためである。NLMS アルゴリズムにおけるステップサイズの時間的な変化を図 5.10, 5.11, 5.12 に示す。入力信号は平均電力が 1 になるように正規化されている。従って、LMS アルゴリズムも NLMS アルゴリズムもステップサイズは平均的には μ_0/N , α/N となる。しかし、 $N = 20$ サンプルにおける $\|\mathbf{x}(n)\|^2$ はこれらの回りで変動する。LMS アルゴリズムでは $\mu = \mu_0/20$ の一定値であるが、NLMS アルゴリズムでは図 5.10, 5.11, 5.12 に示すように 20 サンプル区間の $\|\mathbf{x}(n)\|^2$ によりステップサイズが変動している。

いずれの場合も、ステップサイズの大小によって収束速度が変化することが確認できる。誤差の最小値は $v_2(n)$ の平均電力で決まる。

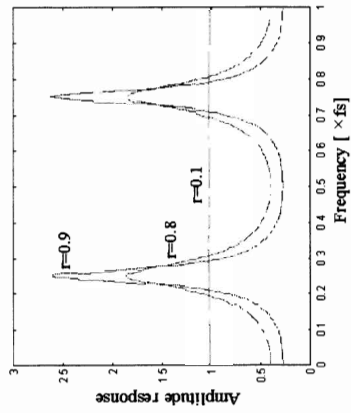


図 5.3: 入力信号を生成する 2 次 AR モデルの振幅特性

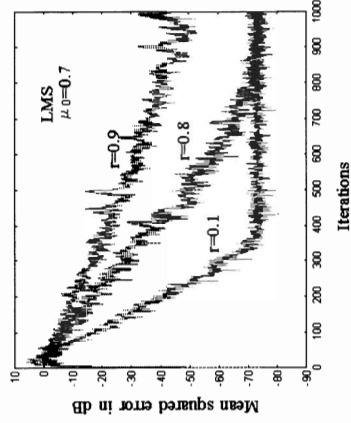


図 5.6: LMS アルゴリズムの学習曲線

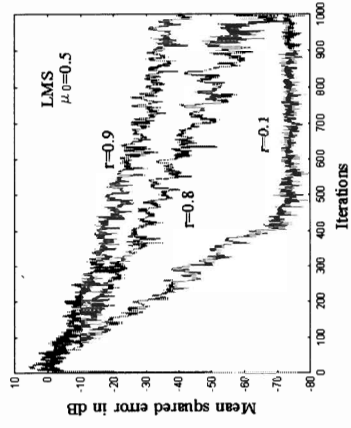


図 5.7: LMS アルゴリズムの学習曲線

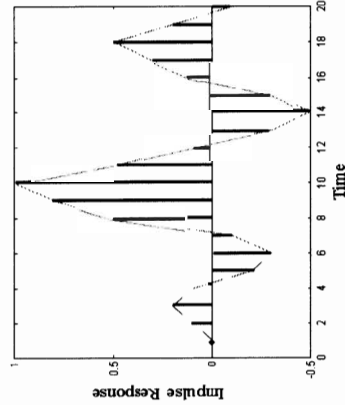


図 5.4: 未知システムのインパルス応答

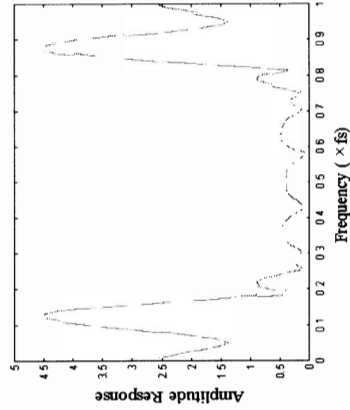


図 5.5: 未知システムの振幅特性

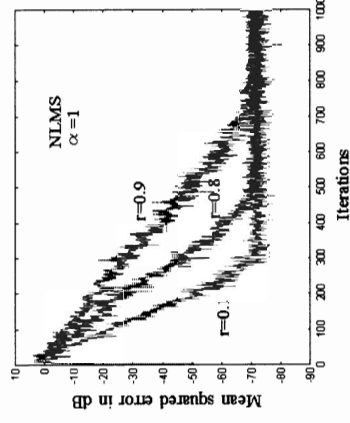


図 5.8: NLMS アルゴリズムの学習曲線

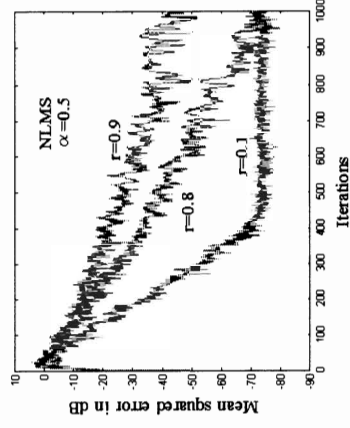


図 5.9: NLMS アルゴリズムの学習曲線

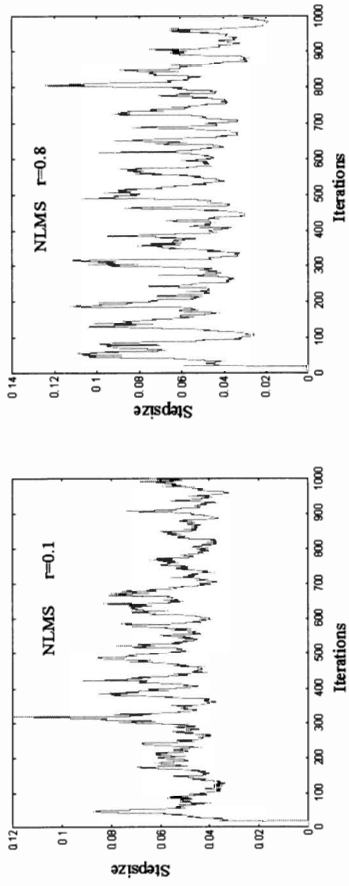


図 5.10: NLMS アルゴリズムのステップサイズ.
図 5.11: NLMS アルゴリズムのステップサイズ.

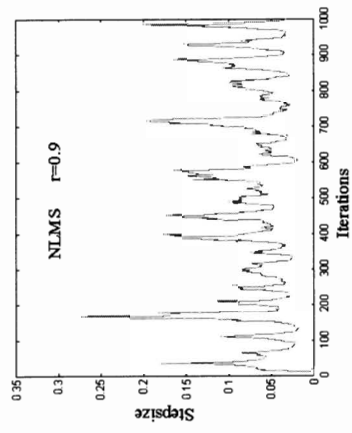


図 5.12: NLMS アルゴリズムのステップサイズ. LMS アルゴリズムのステップサイズは一定値 $\mu = \mu_0/20, \mu_0 = 0.5, 0.7$ である.