ディジタル信号処理

第 10 回

適応信号処理

適応信号処理とは

- 第 8 回で述べたフィルタは, 事前にそれを利用したい状況に関する十分な情報収集をおこなった上で, 何らかの意味で最適に設計されることが通例であるが …
- リアルタイムで情報収集しながらフィルタのパラメータなどを決定したい, という状況もあり得る.

- このような目的で作られる時変フィルタを適応フィルタと呼び, これに対応する信号処理を適応信号処理と呼ぶ.
- 適応信号処理の応用としてはアクティブ消音システムが有名.
- 適応フィルタは時変システムであり, 今まで議論してきた線形時不変システムとは異なった性質を持つ.

- 適応フィルタは, 非再帰型適応フィルタ (**FIR** 型適応フィルタ, トランスバーサルフィルタとも呼ばれる), 再帰型適応フィルタ (**IIR** 型適応フィルタとも呼ばれる), 格子型適応フィルタなどに分類されるが, この講義ではもっとも基本的な非再帰型適応フィルタに議論を限定する.

- 適応信号処理は, 時間領域の適応信号処理と周波数領域の適応信号処理に大別される．この講義では時間領域の適応信号処理のみを取り扱う.
- 適応フィルタは非線形要素を含むこともあるが, この講義では非線形フィルタは取り扱わない.

- x を実数値を取る定常確率過程とする.
- x が実数値で定常だから, その自己相関関数は $\phi[k] = E[x[n]x[n-k]] = E[x[n-k]x[n]] = \phi[-k]$ となる (n は任意).
- $\phi[0] = E[x[l]^2]$ は信号のエネルギーのようなものに相当する.

(教科書では電力と呼んでいる.)

- 教科書では, 112 ページから 113 ページで無限次元フィルタの最適性条件が議論されているが, この講義では, 無限次元フィルタについては議論しない.

非再帰型適応フィルタ

- 以下では，長さ M の非再帰型適応フィルタを検討と対象とする．このフィルタで不規則信号 x を処理するときには，$x[n]$, $x[n-1]$, $\ldots$,$x[n-M+1]$ にフィルタの係数を乗じたものが時刻 n におけるフィルタの出力となる．

非再帰型型適応フィルタ

- $\boldsymbol{x}[n] = (x[n], \ldots, x[n-M+1])^T$ と書く.
- $\boldsymbol{R} = E\left[\boldsymbol{x}[n](\boldsymbol{x}[n])^T\right]$ と定義する. $\boldsymbol{R}$ の要素を具体的に書き下すと, 次のページのようになる.

教科書と比較するときには $\phi[-k] = \phi[k]$ に注意. 教科書では, 106 ページの (9.44) 式と 110 ページの (10.6) 式の整合性が取れていない.

$$
\boldsymbol{R} = \begin{pmatrix} \phi[0] & \phi[1] & & \phi[M-1] \\ \phi[1] & \phi[0] & \ddots & \\ & \ddots & \ddots & \phi[1] \\ \phi[M-1] & & \phi[1] & \phi[0] \end{pmatrix}
$$

- 任意のベクトル $\boldsymbol{v}$ に対し,
  $\boldsymbol{v}^T \boldsymbol{R} \boldsymbol{v} = \boldsymbol{v} E\left[\boldsymbol{x}[n](\boldsymbol{x}[n])^T\right] \boldsymbol{v} = E\left[\left(\boldsymbol{v}^T(\boldsymbol{x}[n])\right)^2\right] \geq 0$
  だから, $\boldsymbol{R}$ は半正定行列である.

信号 x の定常性を仮定すると, 集合平均を標本信号に関する時間平均で置き換えることができ, 教科書で述べられている形が得られる.

$$\boldsymbol{R} = \begin{pmatrix} r[0] & r[1] & & r[M-1] \\ r[1] & r[0] & \ddots & \\ & \ddots & \ddots & r[1] \\ r[M-1] & & r[1] & r[0] \end{pmatrix}$$

- 先に述べた議論をよび前回の議論により, 行列 $\boldsymbol{R}$ は対称なテプリッツ行列で, 半正定値である.
- 以下では, 行列 $\boldsymbol{R}$ は正定値と仮定する. $\{\boldsymbol{x}[n] : n \in \mathbb{Z}\}$ があるベクトル $\boldsymbol{r}$ につねに直交している可能性は非常に低いと考えられるので, 応用上はこの仮定は問題にならない.

- 長さ M の非再帰型フィルタの入出力関係 (x が入力, y が出力) が, $y[n] = \sum_{k=0}^{M-1} w_k[n]x[n-k]$ によって与えられているものとする. $\boldsymbol{w}[n] = (w_0[n], \ldots, w_{M-1}[n])^T$ とおく. $\boldsymbol{w}[n]$ は時変のベクトルであるが, これをある値で固定したものを $\boldsymbol{w}$ と書く.

- フィルタに対する所望信号 d が与えられているとき，$\boldsymbol{x}[n]$ が確定的信号なら，$\boldsymbol{w}^T\boldsymbol{x}[n] = d[n]$ となる $\boldsymbol{w}$ が「望ましいフィルタの係数」であるが，$\boldsymbol{x}$ が不規則信号のときにはこの等号が成立することは期待できない．よって，代替的に，$e[n] = d[n] - y[n]$ とし，次の評価関数 $J(\boldsymbol{w})$ を最小にすることを考える．

$$J(\boldsymbol{w}) = E\left[(e[n])^2\right] = E\left[\left(d[n] - \boldsymbol{w}^T\boldsymbol{x}[n]\right)^2\right]$$

$$
\begin{aligned}
J(\boldsymbol{w}) &= E\left[(d[n])^2 - 2\boldsymbol{w}^T\left(d[n]\boldsymbol{x}[n]\right) + \boldsymbol{w}^T\boldsymbol{x}\boldsymbol{x}^T\boldsymbol{w}\right] \\
&= E\left[(d[n])^2\right] - 2\boldsymbol{w}^T E\left[d[n]\boldsymbol{x}[n]\right] + E\left[\boldsymbol{w}^T\boldsymbol{x}[n](\boldsymbol{x}[n])^T\boldsymbol{w}\right]
\end{aligned}
$$

だから, $\boldsymbol{p} = E\left[d[n]\boldsymbol{x}[n]\right]$, $c = E\left[(d[n])^2\right]$ とおくと,

$$
J(\boldsymbol{w}) = c - 2\boldsymbol{w}^T\boldsymbol{p} + \boldsymbol{w}^T\boldsymbol{R}\boldsymbol{w}.
$$

多くの場合, $\boldsymbol{R}$ は正定対称行列と仮定できるから, $J(\boldsymbol{w})$ は正値二次形式であり, よってある一点 $\boldsymbol{w}_*$ で最適値を取り, かつ $J(\boldsymbol{w}) =$ 一定 という等高線はパラメータベクトルの空間における楕円となる.

- $\boldsymbol{w}$ の最適値 $\boldsymbol{w}_*$ を求めるためには, $\dfrac{\partial J(\boldsymbol{w})}{\partial \boldsymbol{w}} = \boldsymbol{0}$ を $\boldsymbol{w}$ について解けばよい. $J(\boldsymbol{w})$ が二次形式だから, これは以下の連立一次方程式になる:

$$\boldsymbol{p} = \boldsymbol{R}\boldsymbol{w}$$

これをウィーナー・ホップの方程式と呼ぶ. また, その解 $\boldsymbol{w}_*$ をウィーナー解と呼ぶ.

無限次元の場合にも同様に方程式を導くことができ, これもウィーナー・ホップの方程式と呼ぶ. 教科書 113 ページにはこちらが記載されている.

- 行列 $\boldsymbol{R}$ および ベクトル $\boldsymbol{p}$ が既知であれば, ウィーナー・ホップの方程式は連立一次方程式なので, これを解くことは容易.
- 一方, $\boldsymbol{R}$ および $\boldsymbol{p}$ は集合平均 (あるいは無限長の時間平均) によって得られるので, これらの厳密な値は一般には利用不能. 有限長の見本信号の時間平均によってこれらを近似した場合にも, 計算量が多くなる.

- そこで, 1 回の解 (の候補) の更新あたりの計算量が少ない逐次解法で $\boldsymbol{w}$ の最適値を近似的に求めることを考える.
- 以下の議論では, 適応フィルタの分野の標準的かつ古典的な解法である最急降下法と LMS アルゴリズムについて述べる.

最急降下法

- 最急降下法は, 変数ベクトル $\boldsymbol{w}$ の関数 $f(\boldsymbol{w})$ の極小値を求めるための古典的な数値解法(古すぎていつごろ提案されたものかは不明). ディジタル信号処理に固有の方法というわけではなく, 関数の極小値を求める問題で広く利用されている解法である.

- 最急降下法は, 関数 f の勾配ベクトル $\dfrac{\partial f}{\partial \boldsymbol{w}}$ を転置したものが, c を定数としたとき, $\{\boldsymbol{w} : f(\boldsymbol{w}) = c\}$ (関数 f の「等高線」) の外向き法線ベクトルを与えるという性質を利用して, 近似解 $\boldsymbol{w}[k]$ を, 次ページの式によって更新するという解法である.

最急降下法のパラメータ更新則

$$\boldsymbol{w}[k+1] = \boldsymbol{w}[k] - \mu[k] \left(\frac{\partial f}{\partial \boldsymbol{w}} \right)^T \Bigg|_{\boldsymbol{w}=\boldsymbol{w}[k]}$$

- 添字 k は繰り返し計算の回数である.
- パラメータ $\mu[k]$ は, **ステップ幅, ステップサイズパラメータ**などと呼ばれる.

- 最急降下法は, 関数 f の性質が良く, パラメータ $\mu[k]$ が適切であれば, 極小値を与える $\boldsymbol{w}_*$ に収束する列を生成する (この講義では詳細は述べない). 簡単に使えるが収束は遅い. 一般的な最適化では $\mu[k]$ を可変にして収束の遅さを緩和することが普通だが, 適応フィルタでは $\mu[k]$ は多くの場合「十分小さい値」に固定される.

- 適応フィルタにおける最急降下法は, 2 次形式 $J(\boldsymbol{w}) = c - 2\boldsymbol{w}^T\boldsymbol{p} + \boldsymbol{w}^T\boldsymbol{R}\boldsymbol{w}$ の最小化問題に最急降下法を適用したもの.
- 一般的な最適化という観点からは, 正定値の 2 次形式の最小化問題 (連立一次方程式の求解に帰着され, 高速かつ高精度に解ける) に最急降下法を適用することは, 極めて効率が悪い, 「あり得ない」選択肢なのだが …

- 適応フィルタの問題には, 行列 $\boldsymbol{R}$ の厳密な値を求めることが困難であるという特殊な事情があり, このため, 最急降下法を適用することで行列 $\boldsymbol{R}$ の利用を回避することが一般的.

- 適応フィルタにおける最急降下法では,「最適な」パラメータ $\boldsymbol{w}_*$ (ウィーナー・ホップの方程式の解が想定されている) の推定値 $\boldsymbol{w}[k]$ が1サンプリング周期に1回更新されることが一般的. そこで, 以下では, 信号 $\boldsymbol{x}[n]$, パラメータ推定値 $\boldsymbol{w}[n]$ のような書き方をする.

- $J = E\left[(e[n])^2\right]$ とすると，
$$\frac{\partial J}{\partial \boldsymbol{w}} = 2E\left[e[n]\frac{\partial e}{\partial \boldsymbol{w}}\right] = -2E\left[e[n]\left(\boldsymbol{x}[n]\right)^T\right].$$
- 先に述べた最急降下法の式で，$\mu[k] = \frac{1}{2}\mu$ (μ は定数) とすると，

$$\boldsymbol{w}[n+1] = \boldsymbol{w}[n] + \mu E\left[e[n]\boldsymbol{x}[n]\right].$$

- $\mu[k] = (1/2)\mu$ のように 1/2 を乗じたのは, 後の式を見やすくするため,
- 2 次形式の最小化問題をこの形の最急降下法で解いたとき, 近似解が最適値に収束するための十分条件は, $0 < \mu < \dfrac{2}{\lambda_{\max}}$ である. $\lambda_{\max}$ は行列 $\boldsymbol{R}$ の最大固有値で, $\boldsymbol{R}$ が正定対称行列だから実数である. μ が大きいほど収束が早くなるとは限らない (振動的になり得る).

LMS 法

- 今までの議論で, 適応フィルタのための最急降下法において, 見掛け上は行列 $\boldsymbol{R}$ を必要としない形の式が得られたが $\cdots$
- 実はこれは使い勝手が悪い.
- $\boldsymbol{w}[n+1] = \boldsymbol{w}[n] + \mu E\left[e[n]\boldsymbol{x}[n]\right]$ という式をよく見ると $\cdots$

- $E\left[e[n]\boldsymbol{x}[n]\right]$ は集合平均だから, このままの形では使えないし, $e[n] = y[n] - \boldsymbol{w}[n]\boldsymbol{x}[n]$ において, $\boldsymbol{w}[n]$ は時間とともに変動するから, 集合平均を時間平均で置き換えることも厳密には正当化できない (時間平均による近似的な実装は可能).

- 問題となる $E\left[e[n]\boldsymbol{x}[n]\right] = E\left[\boldsymbol{x}[n]e[n]\right]$ に $e[n] = d[n] - \left(\boldsymbol{x}[n]\right)^T \boldsymbol{w}[n]$ を代入して集合平均を取り直し, $\boldsymbol{p} = E\left[d[n]\boldsymbol{x}[n]\right]$, $\boldsymbol{R} = E\left[\boldsymbol{x}[\boldsymbol{n}]\left(\boldsymbol{x}[n]\right)^T\right]$ を思い出すと, 以下のようにはなるが, 結局行列 $\boldsymbol{R}$ が復活してしまう. これを何とかしたい.

$$
\begin{aligned}
E\left[e[n]\boldsymbol{x}[n]\right] &= E\left[\boldsymbol{x}[n]\left(d[n] - \left(\boldsymbol{x}[n]\right)^T \boldsymbol{w}[n]\right)\right] \\
&= \boldsymbol{p} - \boldsymbol{R}\boldsymbol{w}[n]
\end{aligned}
$$

- 適応フィルタのための最急降下法において，集合平均の項を排除し，パラメータ $\boldsymbol{w}[n]$ の更新則を

$$\boldsymbol{w}[n+1] = \boldsymbol{w}[n] + \mu e[n]\boldsymbol{x}[n]$$

に変更したのが，最も基本的な形の **LMS 法** (Widrow(1959)).

B. Widraw, Think about thinking: the discovery of the LMS algorithm, IEEE Signal Processing Magazine, Vol. 22, No. 1, pp. 100–106, 2005 に LMS 法の発見の経緯が述べられている.

- LMS 法は, 収束は遅いものの, 有用なアルゴリズムで, 広く使われており, 今日では様々なバリエーションが知られている.
- 理論的には, 1996 年に, LMS 法が H_∞ の意味で最適であることが示された.
  B. Hassibi, A. H. Sayed and T. Kailath, H^∞ optimality of the LMS algorithm, IEEE Transactions on Signal Processing, Vol. 44, No. 2, pp. 267–280, 1996.

- $\boldsymbol{w}[n]$ の集合平均を取ると,

$$E[\boldsymbol{w}[n+1]] = E[\boldsymbol{w}[n]] + \mu E[e[n]\boldsymbol{x}[n]]$$

となるから, LMS 法は, 平均的には, 最急降下法として動作することになる. よって, ステップ幅 μ を適切に設定すれば, $E[\boldsymbol{w}[n]]$ は最適値 $\boldsymbol{w}_*$ に収束することが期待できる. ただし, この解析法では, $\boldsymbol{w}[n]$ の期待値に関する挙動しかわからない.

- より詳しいLMS法の特性は, 色々な条件を付けた上で分析されるが, 課せられる条件の中に近似的にしか成立しないものがあることと, 解析が繁雑なので, この講義では, 谷荻, カルマンフィルタと適応信号処理, コロナ社, 2005と西山, 最適フィルタリング, 培風館, 2001に従い, 結果のみ紹介する.

- $J(\boldsymbol{w}) = E\left[(d[n])^2\right] - 2\boldsymbol{w}^T\boldsymbol{p} + \boldsymbol{w}^T\boldsymbol{R}\boldsymbol{w}$ であったが, $\boldsymbol{w}$ にウィーナー解 $\boldsymbol{w}_* = \boldsymbol{R}^{-1}\boldsymbol{p}$ を代入したときの最適値を J_* と書くと,

$$J_* = E\left[(d[n])^2\right] - 2\boldsymbol{p}^T\boldsymbol{R}^{-1}\boldsymbol{p}$$

であり, 更に

$$J|_{\boldsymbol{w}=\boldsymbol{w}[n]} = J_* + (\boldsymbol{w}[n] - \boldsymbol{w}_*)^T\boldsymbol{R}(\boldsymbol{w}[n] - \boldsymbol{w}_*)$$

となる.

- LMS 法では, $-2e[n]\boldsymbol{x}[n]$ を $(\partial J/\partial \boldsymbol{w})^T$ の推定値として用いている. この推定の誤差:

$$\boldsymbol{g}[n] = -2e[n]\boldsymbol{x}[n] - (\partial J/\partial \boldsymbol{w})^T$$

  を, 勾配雑音と呼ぶ.

- $\boldsymbol{w}[n] = \boldsymbol{w}_*$ のときは $(\partial J/\partial \boldsymbol{w})\,|_{\boldsymbol{w}=\boldsymbol{w}_*} = \boldsymbol{0}$ なので, 勾配雑音は $-2e[n]\boldsymbol{x}[n]$ となり, その期待値は $(\partial J/\partial \boldsymbol{w})|_{\boldsymbol{w}=\boldsymbol{w}_*} = \boldsymbol{0}$ と一致するから零である.

- $E[\boldsymbol{g}[n]\,(\boldsymbol{g}[n])^T] \simeq 4\boldsymbol{R}J_*$ であることが示される ($\simeq$ は「近似的に等しい」の意味).
- $\boldsymbol{R}$ は対称行列だから, ある直交行列 $\boldsymbol{Q}$ によって対角化される ($\boldsymbol{Q}^T\boldsymbol{R}\boldsymbol{Q} = \mathrm{diag}\{\lambda_1, \ldots, \lambda_M\}$). $\boldsymbol{v}[n] = \boldsymbol{w}[n] - \boldsymbol{w}_*$ とし, $\boldsymbol{Q}\boldsymbol{v}[n] = \boldsymbol{\nu}[n]$ とおく.

- 定常状態では, $k = 1, \ldots, M$ に対し, $\mu\lambda_k < 1$ なら, 近似的に, $E\left[(\nu_k[n])^2\right] = (\mu/(2 - \mu\lambda_i))J_*$ となり, $k \neq l$ なら $E\left[\nu_k[n]\nu_l[n]\right] = 0$ となる.
- この解析がすべての k について可能であるためには, $\mu < 1/\lambda_{\max}$ であればよいが, この条件は LMS 法の収束に関する条件 $(\mu < 2/\lambda_{\max})$ より厳しいことに注意.

- 先の式から, 収束を早めるために μ を大きくするとパラメータ推定値の分散が大きくなることがわかる. μ を小さくすれば分散も小さくなるが収束は遅くなる
- さらに, $\Delta[n] = J|_{\boldsymbol{w}=\boldsymbol{w}[n]} - J_*$ とおくと,

$$E[\Delta[n]] \simeq \frac{1}{2}\mu \operatorname{tr} \boldsymbol{R} J_*$$

となる. これを過剰平均二乗誤差と呼ぶ.

- 先の近似は, $\mu\lambda_{\mathrm{max}}$ が 1 より遥かに小さいことが前提なので注意.
- 過剰平均二乗誤差を J_* で正規化したものを誤調整と呼び, 記号 M であらわす.

$$M = \frac{E[\Delta[n]]}{J_*}$$

である.

- とくに, 信号 x が白色雑音のときには, $E\left[(x[n])^2\right] = \sigma_x^2$ と書くと, M は次のようになる.

$$M \simeq \frac{1}{2} M \mu \sigma_x^2$$

- 以上によって, μ が大きい, あるいは σ_x^2(信号のエネルギーに相当) が大きいほど, 定常状態で得られる評価関数 J の期待値は最適値 J_* より大きくなることがわかる.

正規化 LMS 法

- LMS 法では, $\boldsymbol{x}$ のノルムが小さいほど収束が遅くなる傾向がある．また, $\boldsymbol{x}$ のノルムの期待値が大きい場合, 収束条件を満たすために μ を非常に小さくしなければならない．これらを解消するために考案されたのが**正規化 LMS 法**である．
- 正規化 LMS 法には, 次の 2 種類がある

$$\text{(I)} \quad \boldsymbol{w}[n+1] = \boldsymbol{w}[n] + \mu e[n] \frac{\boldsymbol{x}[n]}{\|\boldsymbol{x}[n]\|}$$

この形のアルゴリズムでは, $\|\boldsymbol{x}[n]\|$ が小さいときには, 零による除算や桁落ちを防ぐためにパラメータ調整機構を停止することが前提である.

$$
\text{(II)} \quad \boldsymbol{w}[n+1] = \boldsymbol{w}[n] + \mu e[n] \frac{\boldsymbol{x}[n]}{\epsilon + \|\boldsymbol{x}[n]\|}
$$

この形のアルゴリズムでは, ϵ は十分小さい正の定数である．この形は **ϵ-LMS 法**と呼ばれることもある (A. H. Sayed, Fundamentals of Adaptive Filtering, Wiley, 2003).

- 正規化 LMS 法では, LMS 法と異なり, ステップ幅 μ に関する条件は

$$0 < \mu < 2$$

であり, 一定の条件が満たされるときには $\mu = 1$ とすることで最大収束速度が得られる.

- 実用上は, 正規化 LMS 法は, 多くの場合 LMS 法より高速であるが, 常にそういうわけではない.

LMS 法のバリエーション

- LMS 法は古典的でかつ有用なアルゴリズムであり, 今日では様々なバリエーションが開発されている.
- この講義では, それらの名前を挙げるのみに留め, 詳細は述べない.

複素 LMS 法 複素信号を取り扱う.

リーキー LMS 法 フィルタ係数に時間とともに減衰する効果を追加する.

モーメンタム LMS 法 フィルタ係数が急激に変化しないようにする.

LMF 法 誤差の評価に e^4 を用いる.

LMS+F 法, LFM/F 法 LMS 法と LMF 法を併用する.

ブロック LMS 法 複数の時刻の信号をまとめて処理する.

変換領域 LMS 法 入力を直交変換することで高速化する.

可変ステップ LMS 法 ステップ幅を可変にする.

適応フィルタの応用

- 適応フィルタは,「多様な分野に応用されている」という程ではないが, 必ずこれに使われるという, いくつかの定番的な応用がある.
- 適応等化器, エコーキャンセラ, アクティブ消音システムの 3 種類の応用問題に, 適応フィルタがよく使われる.

適応等化器

- 通信をおこなう際には, 受信側が受け取る信号は送信信号が符号間干渉や雑音の影響で劣化したものであることが普通.
- 受信信号から送信信号を復元するフィルタのことを等化器と呼ぶ.

- 等化器は, パラメータの決定のためにトレーニングデータを必要とする自動等化器と, それを必要としない適応等化器に大別される.
- 伝送路の特性をリアルタイムで推定して補償する適応等化器は, 適応フィルタの一種になっている. 具体的な構成は複雑なので, この講義では述べない.

エコーキャンセラ

- 電話で会話をする際に, 自分の声が遅延して受話器等から聞こえてくることがある. これをエコーと呼ぶ.

- エコーが発生する経路を推定すれば, 自分の話し声からどのようなエコーが発生するかを予想することができ, 受信音声から予測されるエコーを減算すれば, エコーが少ない受信音声が得られる. このような機構が適応エコーキャンセラである.
- 適応エコーキャンセラは, 構造的には, 標準的な適応フィルタである.

アクティブ消音システム

- 聞き手の位置 (典型的には鼓膜) に入る雑音を推定し, スピーカから雑音と逆位相の音を出せば, 聞き手の位置では雑音が消える．これがアクティブ消音システムで, 典型的な適応フィルタである．ノイズキャンセリングヘッドフォンや車載用音響システムなどで商品化されている．