知的情報処理論 第10回

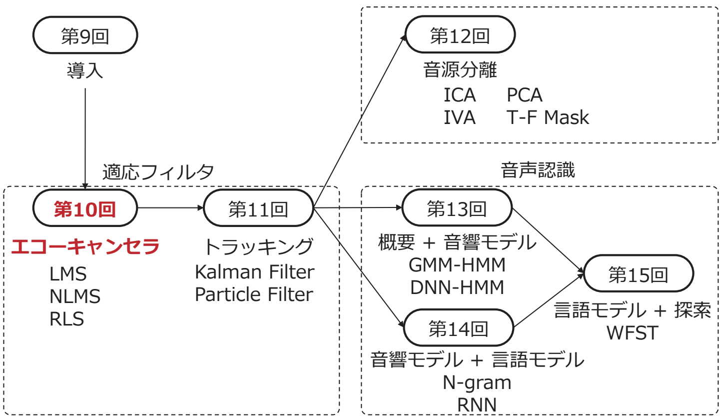
2023年6月20日（火）

武田

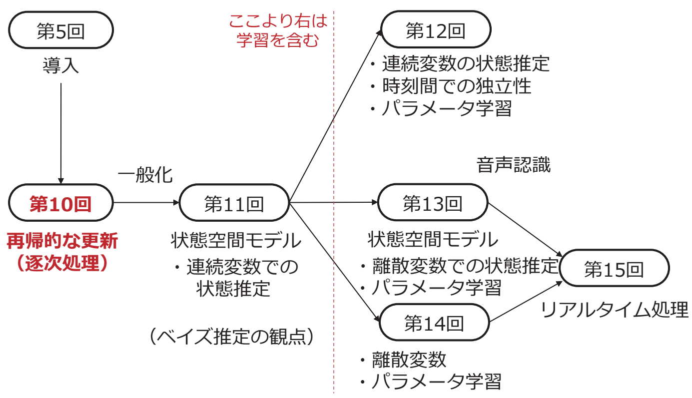
第3回レポート

- 7/3（月） 23:59 提出締め切り
- 詳細は CLE 上の pdf を見ること

各回の内容（予定）



各回の内容（予定）



頭の片隅に置いとく方が良い 大まかな要素

コスト関数
- 「良さ」を測る関数
  - 教師あり学習: 二乗誤差, CrossEntropy, etc…
  - 教師なし学習: 尤度, 再構成誤差, etc…
  - 認識・推定: 期待値, 最大事後確率(MAP), etc…
- ロス/目的関数ともいう

モデル
- 入出力の関係や拘束条件を記述
- 方程式や確率モデルで表現
  （コスト関数と一体化 or 切り離せない場合も）

ざっくりとした 認識/学習/探索
- コスト関数を{最大化|最小化|良く}する値（集合）を求める手続き（最適化など）
- モデルの構造や補助変数・関数を利用した手続きも存在

本日の内容: 適応フィルタ

「再帰的なパラメータ更新へ」

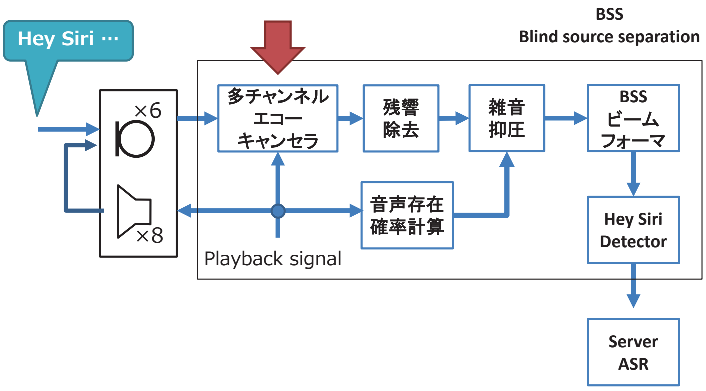
1. ウィーナーフィルタ
2. LMS/NLMS
3. RLS
4. その他

↓ 逐次処理 観測データ利用

↓ 再帰的更新 全観測データ利用

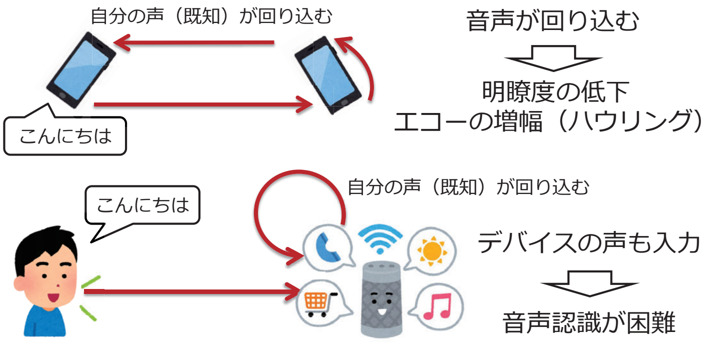
7

Apple Home Pod の構成 (SLT2018より)



8

エコーキャンセラ (適応フィルタ)

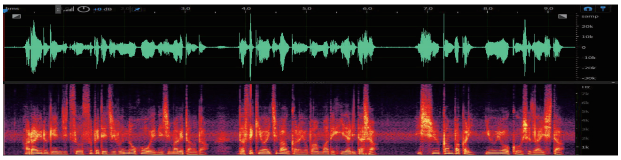


エコー（既知音）を抑圧したい

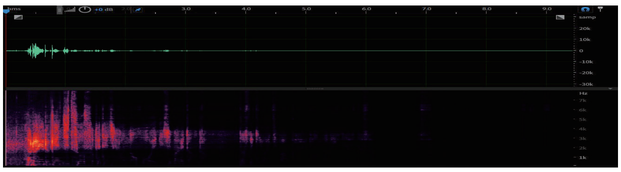
9

エコーキャンセラ

キャンセル前（観測信号）



逐次的なキャンセル



10

動的状況(次回)：トラッキング 斜方投射 + 測定雑音

11

動的状況(次回)：音源定位 (各瞬間瞬間での方向推定)

12

適応フィルタ

具体的な例

目的: 入力からスピーカの音を動的に消す

入力: マイク録音信号 $d[t]$

前提: スピーカから出る音 $u[t]$ は既知（参照信号）

具体的な例

目的: 入力からスピーカの音を動的に消す

入力: マイク録音信号 $d[t]$

前提: スピーカから出る音 $u[t]$ は既知（参照信号）

伝搬経路のモデル

$$y[t]=\sum_{k=0}^{M-1} w[k]u[t-k] \leftrightarrow d[t]$$

同じにしたい

$w[k]$ インパルス応答（パラメータ）

パラメータが求まればスピーカ音の消去が可能

具体的な例

目的: 入力からスピーカの音を動的に消す

入力: マイク録音信号 $d[t]$

前提: スピーカから出る音 $u[t]$ は既知（参照信号）

伝搬経路のモデル

$$y[t]=\sum_{k=0}^{M-1} w[k]u[t-k] \leftrightarrow d[t]$$

同じにしたい

$w[k]$ インパルス応答（パラメータ）

以降の対象

離散時間・線形・有限長のフィルタ（FIR） finite impulse response

パラメータが求まればスピーカ音の消去が可能

フィルタと適応

フィルタ

観測データに対して統計的な処理を行い，目的の信号を抽出する/対象としている量についての正確な情報を抽出する

適応

調整可能なパラメータを持ち，（扱っている信号の）推定した統計的性質に基づいて，それらの値が自動的に指定される

線形・非線形だとかは関係ない

S. Haykin 「適応フィルタ入門」より

基本的な処理

3つの基本的な情報処理操作

- **フィルタリング**
- 平滑化（スムージング）
- 予測

ここでは，時刻 t における

観測データ $\mathbf{x}_t$

適応/推定対象 $\mathbf{z}_t$

次元数は適当に設定されているとする

（潜在変数 or パラメータといった “知りたい” 値）

と表記して説明

フィルタリング（濾波）

時刻 t までに観測したデータを用いて，時刻 t における所望の信号/情報を抽出

対象の変数 $\mathbf{z}_1$ … $\mathbf{z}_{t-2}$ $\mathbf{z}_{t-1}$ $\mathbf{z}_t$ $\mathbf{z}_{t+1}$

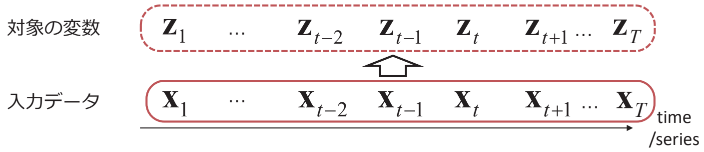
入力データ $\mathbf{x}_1$ … $\mathbf{x}_{t-2}$ $\mathbf{x}_{t-1}$ $\mathbf{x}_t$ $\mathbf{x}_{t+1}$

time /series

- 逐次処理（リアルタイム処理向き）
- 密度/分布関数で捉えた場合 $p(\underbrace{\mathbf{z}_t}_{\text{知りたい値}} \mid \underbrace{\mathbf{x}_1,\ldots,\mathbf{x}_t}_{\text{与えられたデータ}})$

平滑化（スムージング）

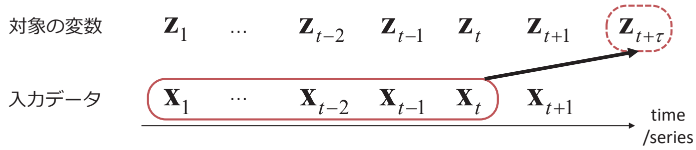
時刻 t 以降に観測したデータも用いて，所望の信号/情報を抽出



– 未来のデータも利用可能（バッチ処理）
　= より正確な抽出が期待
– 過去のデータを処理する場合は遅延が生じる
– 密度/分布関数で捉えた場合 $p(\mathbf{z}_1,\ldots,\mathbf{z}_T \mid \mathbf{x}_1,\ldots\mathbf{x}_T)$

予測

時刻 t までの観測データを用いて，ある $\tau > 0$ だけ先の時刻 $t+\tau$ の所望の信号/情報がどのようになるかを求める



– 時間発展を考慮
– 密度/分布関数で捉えた場合 $p(\mathbf{z}_{t+\tau} \mid \mathbf{x}_1,\ldots,\mathbf{x}_t)$

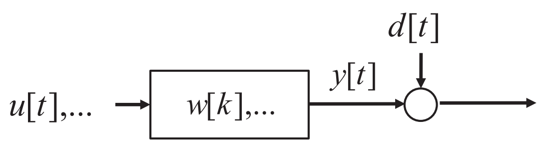
適応信号処理の例

未知の動的システムのモデリング
– System Identification

音声生成過程のモデル
– 声道フィルタ: 音声特徴量と関連

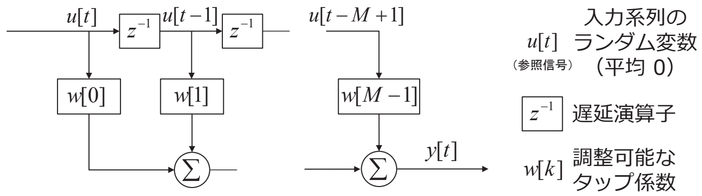
線形予測符号化
– linear prediction による音声パラメータ推定

エコーキャンセラ・適応ビームフォーマ



ウィーナーフィルタ

ブロック図と問題設定



目的: 望みの応答 $d[t]$ と出力 $y[t]$ の差の最小化

$y[t] = \sum_{k=0}^{M-1} w[k]u[t-k]$　　誤差信号 $e[t] = d[t] - y[t]$

ウィーナー理論: フィルタは誤差信号の二乗平均値 $E[|e_t|^2]$ を最小にするように最適化

ベクトル表現に変換
インデックスの表記変更

$$y[t] = \sum_{k=0}^{M-1} w[k]u[t-k]$$

各変数

$u_t = u[t] \quad y_t = y[t] \quad w_k = w[k]$

$d_t = d[t] \quad e_t = d_t - y_t$

入力信号（参照信号）ベクトル

$$\mathbf{u}_t = [u_t, u_{t-1}, \ldots, u_{t-M+1}]^{\mathrm{T}}$$

フィルタ係数ベクトル

$$\mathbf{w} = [w_0, w_1, \ldots, w_{M-1}]^{\mathrm{T}}$$

$$y_t = \mathbf{w}^{\mathrm{T}}\mathbf{u}_t$$

仮定

ランダム時系列 u_t は定常

以下の2つの条件を満たす

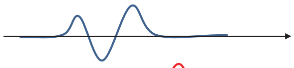
1. この過程の平均値 m は時刻 t に独立

$$m = E[u_t] = \mathrm{const.}$$

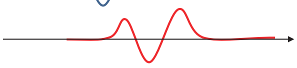
E 期待値演算
以降，m=0 を仮定

2. 自己相関関数は時間差 t – k のみに依存

$$r[t,k] \triangleq E[u_t u_k]$$



$$r[t,k] = r[t-k]$$



定常：信号過程と雑音過程の統計的性質が時間により変化しない
＝ある時刻のデータから推測した情報を，異なる時刻での推測に利用できるような性質

ウィーナーフィルタの導出

誤差の平均二乗誤差をコスト関数として最小化

まず展開する

$\mathbf{w}$ 定数
$d_t, \mathbf{u}_t$ 密度関数の形状や条件には触れてない

$$J = E[|e_t|^2]$$

$$= E[(d_t - \mathbf{w}^{\mathrm{T}}\mathbf{u}_t)(d[t] - \mathbf{w}^{\mathrm{T}}\mathbf{u}_t)^{\mathrm{T}}]$$

$$= E[|d_t|^2] - 2\mathbf{w}^{\mathrm{T}} E[d_t \mathbf{u}_t]$$

$$+ \mathbf{w}^{\mathrm{T}} E[\mathbf{u}_t \mathbf{u}_t^{\mathrm{T}}]\mathbf{w}$$

ウィーナーフィルタの導出

記号を定義して整理

望みの応答の2乗平均値

$$\sigma_d^2 = E[|d_t|^2]$$

望みの応答と入力信号の相互相関ベクトル

$$\mathbf{r} = E[d_t \mathbf{u}_t]$$

結合定常を仮定：d_t と u_{t-k+1} が定常
相互相関関数が時間差 k - 1 に依存

自己相関関数

$$\mathbf{R} = E[\mathbf{u}_t \mathbf{u}_t^{\mathrm{T}}]$$

Toeplitz（テプリッツ）行列
主体対角要素がすべて同じ
並行な他の対角線上の要素も等しい

i 行 j 列目 $E[u_{t-i+1} u_{t-j+1}] = r[j-i]$

補足：自己相関行列の要素

$$\mathbf{R} = E[\mathbf{u}_t \mathbf{u}_t^{\mathrm{T}}] = E[\begin{bmatrix} u_t \\ u_{t-1} \\ \vdots \\ u_{t-M+1} \end{bmatrix} [u_t \quad u_{t-1} \quad \cdots \quad u_{t-M+1}]]$$

$$= \begin{bmatrix} E[u_t u_t] & E[u_t u_{t-1}] & \cdots & E[u_t u_{t-M+1}] \\ E[u_{t-1} u_t] & E[u_{t-1} u_{t-1}] & \cdots & E[u_{t-1} u_{t-M+1}] \\ \vdots & \vdots & \ddots & \vdots \\ E[u_{t-M+1} u_t] & E[u_{t-M+1} u_{t-1}] & \cdots & E[u_{t-M+1} u_{t-M+1}] \end{bmatrix}$$

仮定の2 $r[t,k] = E[u_t u_k] = r[t-k]$

ウィーナーフィルタの導出

定義した記号で置換

$$J = \sigma_d^2 - 2\mathbf{w}^{\mathrm{T}}\mathbf{r} + \mathbf{w}^{\mathrm{T}}\mathbf{R}\mathbf{w}$$

$\min_{\mathbf{w}} J(\mathbf{w})$ を解く問題

コスト関数の導関数が０になる時に最小

$$\frac{\partial J}{\partial \mathbf{w}} = -2\mathbf{r} + 2\mathbf{R}\mathbf{w} = \mathbf{0}$$

$\mathbf{R}\mathbf{w} = \mathbf{r}$ **正規方程式/ウィーナー・ホッフ方程式（離散時間）**

ウィーナーフィルタの導出

最適なフィルタ係数 $\hat{\mathbf{w}}$

相関行列の逆行列をかける

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

自己相関行列はほとんど常に
正定値行列

最小平均二乗誤差

最適値なフィルタ値を代入

$$J_{\min} = \sigma_d^2 - \mathbf{r}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{r}$$

これ以上誤差を小さくする線形フィルタは
設計できない

直交原理

正規方程式 $\mathbf{R}\hat{\mathbf{w}} = \mathbf{r}$ を変形

$$E[\mathbf{u}_t \mathbf{u}_t^{\mathrm{T}}]\hat{\mathbf{w}} = E[d_t \mathbf{u}_t]$$
$$E[(d_t - \hat{\mathbf{w}}^{\mathrm{T}}\mathbf{u}_t)\mathbf{u}_t] = \mathbf{0}$$
$$E[(d_t - \hat{y}_t)\mathbf{u}_t] = \mathbf{0}$$

$$E[\hat{e}_t \mathbf{u}_t] = \mathbf{0}$$

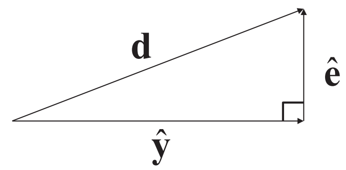
$\hat{e}_t$ 最適フィルタを用いた時に生じる誤差信号

誤差信号とすべてのタップ係数は直交＝直交原理

$$E[\hat{e}_t \hat{y}_t] = E[\hat{e}_t \hat{\mathbf{w}}^{\mathrm{T}}\mathbf{u}_t] = \hat{\mathbf{w}}^{\mathrm{T}} E[\hat{e}_t \mathbf{u}_t] = \mathbf{0}$$

誤差信号と最適フィルタ出力も直交

直交原理の図的解釈



- 各変数をベクトルと”みなした”幾何学的関係
- 誤差ベクトル $\hat{\mathbf{e}}$ とフィルタ出力ベクトル $\hat{\mathbf{y}}$ は直交するように描かれる
- 誤差が最小となるのは直交が満足されている時のみ
  - 誤差は白色

有限サンプルを用いた近似 **最小二乗法**

期待値演算: 確率密度・積分計算が必要
→ データのサンプル（平均）で置き換え

コスト関数（二乗誤差の総和）

$$\hat{J} = \frac{1}{T}\sum_{t=0}^{T-1} |e_t|^2 \qquad (T > M)$$
$$= \frac{1}{T}\sum_{t=0}^{T-1} (d_t - \mathbf{w}^{\mathrm{T}}\mathbf{u}_t)(d_t - \mathbf{w}^{\mathrm{T}}\mathbf{u}_t)^{\mathrm{T}}$$

解く問題 $\min_{\mathbf{w}} \hat{J}(\mathbf{w}) \qquad \hat{\mathbf{w}} = \arg\min_{\mathbf{w}} \hat{J}(\mathbf{w})$

有限サンプルを用いた近似 **最小二乗法**

期待値演算: 確率密度・積分計算が必要
→ データのサンプル（平均）で置き換え

最適フィルタ $\hat{\mathbf{w}} = \hat{\mathbf{R}}^{-1}\hat{\mathbf{r}}$

$$\hat{\mathbf{r}} = \frac{1}{T}\sum_{t=0}^{T-1} d_t \mathbf{u}_t \qquad \hat{\mathbf{R}} = \frac{1}{T}\sum_{t=0}^{T-1} \mathbf{u}_t \mathbf{u}_t^{\mathrm{T}}$$

$$\hat{\sigma}_d^2 = \frac{1}{T}\sum_{t=0}^{T-1} |d_t|^2$$

Mini Quiz #1

- 入力信号 u_t が白色雑音（平均0, 分散1のガウス分布から生成）の場合，自己相関行列 $\mathbf{R}$はどのような形になるか？

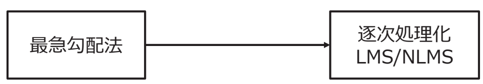
$$u_t \sim N(0,1)$$

$\mathbf{u}_t \quad \mathbf{R}$

- 計算機で確かめる場合: 十分大きなサンプル数をもって，サンプル平均で近似計算

Octave 疑似コード

```
N = 10000;
n = randn(15, N);
R = n * n' / N;
surf(R); % 3Dプロット
```

LMS/NLMS

リアルタイム/逐次処理

エコーキャンセラ
- 動的にフィルタ係数を推定したい
  - 途中で伝達関数が変わっても追従してほしい
- 処理量やメモリ量に制約がある場合も存在
  - 安価・小型デバイス

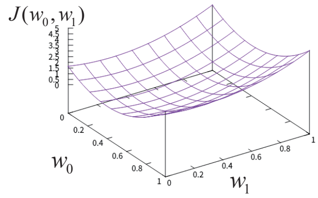
ウィーナーフィルタ・最小二乗法
- 期待値演算 → あまり単純な処理ではない

Least Mean Square (LMS) 法/最小二乗平均法

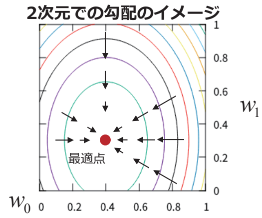
（その前に）最急降下法

正規方程式の解を反復法で求める
- コスト関数 J（2次関数）の曲面：誤差特性曲面
- 勾配を推定し，逆の方向にフィルタ係数を変化

2次元での曲面の例



2次元での勾配のイメージ



（その前に）最急降下法

正規方程式の解を反復法で求める
- コスト関数 J（2次関数）の曲面：誤差特性曲面
- 勾配を推定し，逆の方向にフィルタ係数を変化

$$\nabla J(\mathbf{w}_{k-1}) = \left.\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}}\right|_{\mathbf{w}=\mathbf{w}_{k-1}}$$

k 反復数

$$\mathbf{w}_k = \mathbf{w}_{k-1} - \mu \nabla J(\mathbf{w}_{k-1})$$

μ 学習係数（ステップサイズパラメータ）

（その前に）最急降下法

正規方程式の解を反復法で求める
- コスト関数 J（2次関数）の曲面：誤差特性曲面
- 勾配を推定し，逆の方向にフィルタ係数を変化

$$\nabla J(\mathbf{w}_{k-1}) = \left.\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}}\right|_{\mathbf{w}=\mathbf{w}_{k-1}} = -2\mathbf{r} + 2\mathbf{R}\mathbf{w}_{k-1}$$

具体的な更新式

$$\boxed{\mathbf{w}_k = \mathbf{w}_{k-1} + 2\mu(\mathbf{r} - \mathbf{R}\mathbf{w}_{k-1})}$$

データ全部利用 ＝ バッチ処理

（その前に）最急降下法

更新式を整理

結果がスカラー値なら

$$\mathbf{u}_t^{\mathrm{T}}\mathbf{w}_{k-1} = \mathbf{w}_{k-1}^{\mathrm{T}}\mathbf{u}_t$$
$$\mathbf{u}_t(\mathbf{w}_{k-1}^{\mathrm{T}}\mathbf{u}_t) = (\mathbf{w}_{k-1}^{\mathrm{T}}\mathbf{u}_t)\mathbf{u}_t$$

$$\begin{aligned}
\mathbf{w}_k &= \mathbf{w}_{k-1} + 2\mu(\mathbf{r} - \mathbf{R}\mathbf{w}_{k-1}) \\
&= \mathbf{w}_{k-1} + 2\mu(E[d_t\mathbf{u}_t] - E[\mathbf{u}_t\mathbf{u}_t^{\mathrm{T}}]\mathbf{w}_{k-1}) \\
&= \mathbf{w}_{k-1} + 2\mu E[(d_t - \mathbf{w}_{k-1}^{\mathrm{T}}\mathbf{u}_t)\mathbf{u}_t]
\end{aligned}$$

$$e_t = d_t - \mathbf{w}_{k-1}^{\mathrm{T}}\mathbf{u}_t \Leftarrow$$

現状のパラメータを用いた事前推定誤差（予測誤差/誤差信号）

$$\underline{\mathbf{w}_k = \mathbf{w}_{k-1} + 2\mu E[e_t\mathbf{u}_t]}$$

係数の2は学習係数に含めても良い

LMS法

勾配の瞬時推定値を利用
- 各時刻毎にフィルタの更新が可能
- 1サンプルに対する勾配を利用（確率的勾配法）

最急降下（バッチ）
$$\mathbf{w}_k = \mathbf{w}_{k-1} + 2\mu E[e_t\mathbf{u}_t] \qquad E[e_t\mathbf{u}_t] \approx \frac{1}{T}\sum_t e_t\mathbf{u}_t$$

LMS（逐次）
$$\mathbf{w}_t = \mathbf{w}_{t-1} + 2\mu e_t\mathbf{u}_t$$

添え字に注意
反復数 → 時間ステップ
全サンプル → 1サンプル

時刻 t のコスト関数
$$J_t = (d_t - \mathbf{w}_{t-1}^T\mathbf{u}_t)^2 \Rightarrow \frac{\partial J_t}{\partial \mathbf{w}_{t-1}} = -2\underbrace{(d_t - \mathbf{w}_{t-1}^T\mathbf{u}_t)}_{e_t}\mathbf{u}_t$$

e_t 誤差信号

NLMS 法

入力のパワーで正規化した LMS

– Normalized LMS

– LMSよりも収束性が良い

$$\mathbf{w}_t = \mathbf{w}_{t-1} + \mu \frac{1}{\|\mathbf{u}_t\|^2}\mathbf{u}_t e_t$$

– 入力信号のパワーが非常に小さい時は不安定

- 小さい値 α を加える: 正則化

$$\mathbf{w}_t = \mathbf{w}_{t-1} + \mu \frac{1}{\alpha + \|\mathbf{u}_t\|^2} e_t \mathbf{u}_t$$

NLMS の定式化

下記の**拘束付き最適化問題から導出**

$$\min_{\mathbf{w}_t} \|\mathbf{w}_t - \mathbf{w}_{t-1}\|^2 \quad \text{subject to} \;\; \mathbf{w}_t^{\mathrm{T}}\mathbf{u}_t = d_t$$

Principle of minimal disturbance

– ラグランジュの未定乗数法を用いて解く

$$J = \|\mathbf{w}_t - \mathbf{w}_{t-1}\|^2 + \lambda(d_t - \mathbf{w}_t^{\mathrm{T}}\mathbf{u}_t)$$

$$= (\mathbf{w}_t - \mathbf{w}_{t-1})^{\mathrm{T}}(\mathbf{w}_t - \mathbf{w}_{t-1}) + \lambda(d_t - \mathbf{w}_t^{\mathrm{T}}\mathbf{u}_t)$$

NLMS の定式化

– $\mathbf{w}_t$ について偏微分

$$\frac{\partial J}{\partial \mathbf{w}_t} = 2(\mathbf{w}_t - \mathbf{w}_{t-1}) - \lambda \mathbf{u}_t \quad \Rightarrow \quad \mathbf{w}_t - \mathbf{w}_{t-1} = \frac{1}{2}\lambda \mathbf{u}_t$$

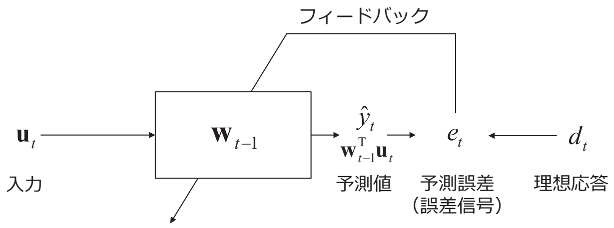
– 左から $\mathbf{u}_t^{\mathrm{T}}$ を乗じる $\Rightarrow$ $\lambda \mathbf{u}_t^{\mathrm{T}}\mathbf{u}_t = 2\mathbf{u}_t^{\mathrm{T}}(\mathbf{w}_t - \mathbf{w}_{t-1})$

$$\lambda = \frac{2}{\mathbf{u}_t^{\mathrm{T}}\mathbf{u}_t}\mathbf{u}_t^{\mathrm{T}}(\mathbf{w}_t - \mathbf{w}_{t-1})$$

$$= \frac{2}{\|\mathbf{u}_t\|^2}(d_t - \mathbf{u}_t^{\mathrm{T}}\mathbf{w}_{t-1}) = \frac{2}{\|\mathbf{u}_t\|^2} e_t$$

$$\mathbf{w}_t = \mathbf{w}_{t-1} + \frac{1}{2}\lambda \mathbf{u}_t = \mathbf{w}_{t-1} + \frac{1}{\|\mathbf{u}_t\|^2} e_t \mathbf{u}_t$$

処理の流れ



LMS $\quad \mathbf{w}_t = \mathbf{w}_{t-1} + 2\mu e_t \mathbf{u}_t$

NLMS $\quad \mathbf{w}_t = \mathbf{w}_{t-1} + \mu \frac{1}{\alpha + \|\mathbf{u}_t\|^2} e_t \mathbf{u}_t$

フィルタリング $\quad y_t = \mathbf{w}_t^{\mathrm{T}}\mathbf{u}_t$

Mini Quiz #2

- ある環境のフィルタ（伝達特性）を LMSで推定したい
- 音量が異なる信号 $u[t]$ と $0.001u[t]$ を入力信号として，フィルタ更新値の違いを比較する
- 学習係数としてある値 μ を用いる場合，2つの入力信号間の音量の違いが，フィルタ更新式のどの部分に現れるだろうか
  – 初期値 $\mathbf{w}_0 = \mathbf{0}$ として，$\mathbf{w}_1$ の違いを考えてみよう

LMS $\quad \mathbf{w}_t = \mathbf{w}_{t-1} + 2\mu e_t \mathbf{u}_t \quad e_t = d_t - \mathbf{w}_{t-1}^{\mathrm{T}}\mathbf{u}_t$

真の伝達関数（固定）$\mathbf{h}$ $\quad d_t = \mathbf{h}^{\mathrm{T}}\mathbf{u}_t$

RLS

再帰最小二乗法（RLS）

Recursive least square

最小二乗解を再帰的に求める手法

– 統計量の仮定無し: 得られた観測データすべてに関して誤差を最小化
– 入力毎にフィルタ係数を更新
  ☺ LMS よりもはるかに速い収束速度
  ☹ 計算量とのトレードオフ
– d_t と $\mathbf{u}_t$ が結合定常[1]なエルゴード過程に従う
  集合平均＝時間平均
  → データ数無限大でウィーナー解に近づく

考え方の参考のために導出を追う

（計算結果の再利用＋逐次更新）

[1] d_t と u_{t-k+1} が定常，相互相関関数が時間差 k - 1 に依存

RSL 定式化

コスト関数: 重み付き誤差の場合

– 過去の誤差の影響を小さくする
  → パラメータ変化に対する追従を重視
– 重み λ $(0<\lambda<1)$ を導入

$$J_t = \sum_{i=0}^{t} \lambda^{t-i} |e_i|^2 = \sum_{i=0}^{t} \lambda^{t-i} |d_i - \mathbf{w}_t^{\mathrm{T}} \mathbf{u}_i|^2$$

RSL 定式化

一旦展開して最適解を求める

$$\begin{aligned} J_t &= \sum_{i=0}^{t} \lambda^{t-i} (d_i - \mathbf{w}_t^{\mathrm{T}} \mathbf{u}_i)(d_i - \mathbf{w}_t^{\mathrm{T}} \mathbf{u}_i)^{\mathrm{T}} \\ &= \sum_{i=0}^{t} \lambda^{t-i} (d_i^2 - 2 d_i \mathbf{w}_t^{\mathrm{T}} \mathbf{u}_i + \mathbf{w}_t^{\mathrm{T}} \mathbf{u}_i \mathbf{u}_i^{\mathrm{T}} \mathbf{w}_t) \\ &= \sum_{i=0}^{t} \lambda^{t-i} d_i^2 - 2\mathbf{w}_t^{\mathrm{T}} \sum_{i=0}^{t} \lambda^{t-i} d_i \mathbf{u}_i \\ &\qquad + \mathbf{w}_t^{\mathrm{T}} \left(\sum_{i=0}^{t} \lambda^{t-i} \mathbf{u}_i \mathbf{u}_i^{\mathrm{T}} \right) \mathbf{w}_t \end{aligned}$$

RSL 定式化

偏微分して 0 と置く

$$\left(\sum_{i=0}^{t} \lambda^{t-i} \mathbf{u}_i \mathbf{u}_i^{\mathrm{T}} \right) \mathbf{w}_t = \sum_{i=0}^{t} \lambda^{t-i} d_i \mathbf{u}_i$$

– 自己相関行列と相互相関ベクトルの推定値

$$\mathbf{R}_t = \left(\sum_{i=0}^{t} \lambda^{t-i} \mathbf{u}_i \mathbf{u}_i^{\mathrm{T}} \right) \qquad \mathbf{r}_t = \sum_{i=0}^{t} \lambda^{t-i} d_i \mathbf{u}_i$$

$$\mathbf{R}_t \mathbf{w}_t = \mathbf{r}_t \qquad \mathbf{w}_t = \mathbf{R}_t^{-1} \mathbf{r}_t$$

これまでと
似たような形

再帰的な表現の導出

方針: $\mathbf{w}_t = \mathbf{R}_t^{-1} \mathbf{r}_t$ を $\mathbf{w}_{t-1}$ や t - 1 時刻の情報で記述

– 次の変数を上手く変形
  自己相関関数

$$\mathbf{R}_t = \left(\sum_{i=0}^{t} \lambda^{t-i} \mathbf{u}_i \mathbf{u}_i^{\mathrm{T}} \right) = \lambda \mathbf{R}_{t-1} + \mathbf{u}_t \mathbf{u}_t^{\mathrm{T}}$$

  相互相関ベクトル

$$\mathbf{r}_t = \sum_{i=0}^{t} \lambda^{t-i} d_i \mathbf{u}_i = \lambda \mathbf{r}_{t-1} + d_t \mathbf{u}_t$$

まずやっかいな逆行列部分を何とかしていく

再帰的な表現の導出

– 逆行列 $\mathbf{R}_t^{-1} = \mathbf{P}_t$ は逆行列の補助定理により

$$\mathbf{R}_t = \lambda \mathbf{R}_{t-1} + \mathbf{u}_t \mathbf{u}_t^{\mathrm{T}}$$

$$\mathbf{A} = \mathbf{B}^{-1} + \mathbf{C}\mathbf{D}^{-1}\mathbf{C}^{\mathrm{T}}$$

$$\mathbf{A}^{-1} = \mathbf{B} - \mathbf{B}\mathbf{C}(\mathbf{D} + \mathbf{C}^{\mathrm{T}}\mathbf{B}\mathbf{C})^{-1}\mathbf{C}^{\mathrm{T}}\mathbf{B}$$

$$\mathbf{P}_t = \lambda^{-1} \underline{\mathbf{P}_{t-1}} - \frac{\lambda^{-2} \mathbf{P}_{t-1} \mathbf{u}_t \mathbf{u}_t^{\mathrm{T}} \mathbf{P}_{t-1}}{1 + \lambda^{-1} \mathbf{u}_t^{\mathrm{T}} \mathbf{P}_{t-1} \mathbf{u}_t}$$

とりあえず受け入れる
（存在は知っておく）

– ゲインベクトル $\mathbf{g}_t$ を定義して整理

$$\mathbf{g}_t = \frac{\lambda^{-1} \mathbf{P}_{t-1} \mathbf{u}_t}{1 + \lambda^{-1} \mathbf{u}_t^{\mathrm{T}} \mathbf{P}_{t-1} \mathbf{u}_t}$$

これも受け入れる

$$\mathbf{P}_t = \lambda^{-1} \mathbf{P}_{t-1} - \lambda^{-1} \mathbf{g}_t \mathbf{u}_t^{\mathrm{T}} \mathbf{P}_{t-1}$$

再帰的な表現の導出

- $\mathbf{P}_t$ の変形

$$\mathbf{P}_t = \lambda^{-1}\mathbf{P}_{t-1} - \lambda^{-1}\mathbf{g}_t\mathbf{u}_t^{\mathrm{T}}\mathbf{P}_{t-1}$$
$$= \lambda^{-1}(\mathbf{I} - \mathbf{g}_t\mathbf{u}_t^{\mathrm{T}})\mathbf{P}_{t-1}$$

- $\mathbf{g}_t$ の変形

$$\mathbf{g}_t(1 + \lambda^{-1}\mathbf{u}_t^{\mathrm{T}}\mathbf{P}_{t-1}\mathbf{u}_t) = \lambda^{-1}\mathbf{P}_{t-1}\mathbf{u}_t$$
$$\mathbf{g}_t = \lambda^{-1}\mathbf{P}_{t-1}\mathbf{u}_t - \mathbf{g}_t\lambda^{-1}\mathbf{u}_t^{\mathrm{T}}\mathbf{P}_{t-1}\mathbf{u}_t$$
$$= \lambda^{-1}(\mathbf{I} - \mathbf{g}_t\mathbf{u}_t^{\mathrm{T}})\mathbf{P}_{t-1}\mathbf{u}_t$$
$$= \mathbf{P}_t\mathbf{u}_t$$

再帰的な表現の導出

- 最適解に戻ってきて変形

$$\mathbf{w}_t = \mathbf{P}_t\mathbf{r}_t$$
$$= \lambda\mathbf{P}_t\mathbf{r}_{t-1} + \mathbf{P}_t\mathbf{u}_t d_t$$
$$= (\mathbf{I} - \mathbf{g}_t\mathbf{u}_t^{\mathrm{T}})\mathbf{P}_{t-1}\mathbf{r}_{t-1} + \mathbf{g}_t d_t$$
$$= (\mathbf{I} - \mathbf{g}_t\mathbf{u}_t^{\mathrm{T}})\underline{\mathbf{w}_{t-1}} + \mathbf{g}_t d_t$$
$$= \mathbf{w}_{t-1} + \mathbf{g}_t(d_t - \mathbf{w}_{t-1}^{\mathrm{T}}\mathbf{u}_t)$$
$$= \underline{\mathbf{w}_{t-1} + \mathbf{g}_t e_t}$$

$$\left\{\ \mathbf{r}_t = \lambda\mathbf{r}_{t-1} + d_t\mathbf{u}_t\right.$$
$$\left\{\begin{array}{l}\mathbf{P}_t = \lambda^{-1}(\mathbf{I} - \mathbf{g}_t\mathbf{u}_t^{\mathrm{T}})\mathbf{P}_{t-1} \\ \mathbf{g}_t = \mathbf{P}_t\mathbf{u}_t\end{array}\right.$$
$$\left\{\ \mathbf{w}_{t-1} = \mathbf{P}_{t-1}\mathbf{r}_{t-1}\right.$$

RLS まとめ

初期化 $\mathbf{P}_0 = \varepsilon\mathbf{I}$ $\quad \mathbf{w}_0 = \mathbf{0}$ $\quad \varepsilon > 0$ 小さい定数 Pの正定値保証

反復

ゲインベクトル $\mathbf{g}_t = \dfrac{\lambda^{-1}\mathbf{P}_{t-1}\mathbf{u}_t}{1 + \lambda^{-1}\mathbf{u}_t^{\mathrm{T}}\mathbf{P}_{t-1}\mathbf{u}_t}$

予測誤差 $e_t = d_t - \mathbf{w}_{t-1}^{\mathrm{T}}\mathbf{u}_t$

フィルタ更新 $\mathbf{w}_t = \mathbf{w}_{t-1} + \mathbf{g}_t e_t$

自己相関行列の逆行列更新 $\mathbf{P}_t = \lambda^{-1}(\mathbf{I} - \mathbf{g}_t\mathbf{u}_t^{\mathrm{T}})\mathbf{P}_{t-1}$

フィルタリング $y_t = \mathbf{w}_t^{\mathrm{T}}\mathbf{u}_t$

LMS と RLS の対比

更新式

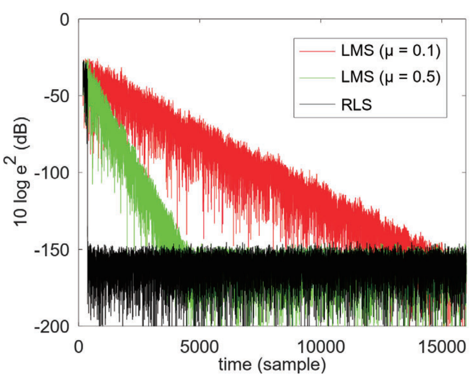
LMS $\mathbf{w}_t = \mathbf{w}_{t-1} + 2\mu e_t\mathbf{u}_t$

RLS $\mathbf{w}_t = \mathbf{w}_{t-1} + \mathbf{g}_t e_t$ $\qquad \mathbf{g}_t = \dfrac{\lambda^{-1}\mathbf{P}_{t-1}\mathbf{u}_t}{1 + \lambda^{-1}\mathbf{u}_t^{\mathrm{T}}\mathbf{P}_{t-1}\mathbf{u}_t}$

$$= \mathbf{w}_{t-1} + \frac{\lambda^{-1}}{1 + \lambda^{-1}\mathbf{u}_t^{\mathrm{T}}\mathbf{P}_{t-1}\mathbf{u}_t} e_t\mathbf{P}_{t-1}\mathbf{u}_t$$

RLS は $\mathbf{P} = \mathbf{R}^{-1}$ による補正
→ 入力の相関をキャンセルする効果
（白色性の信号 vs. 音声信号など（有色））

誤差信号のプロット



再掲: フィルタリング 次回はより抽象化して扱う

対象の変数（重み） $\mathbf{w}_1 \rightarrow \cdots\ \mathbf{w}_{t-2} \rightarrow \mathbf{w}_{t-1} \rightarrow \boxed{\mathbf{w}_t} \rightarrow \mathbf{w}_{t+1}$

データ 観測 $d_1 \ \cdots \ d_{t-2} \ d_{t-1} \ d_t \ d_{t+1}$

データ 内部保持 $\mathbf{u}_1 \ \cdots \ \mathbf{u}_{t-2} \ \mathbf{u}_{t-1} \ \mathbf{u}_t \ \mathbf{u}_{t+1}$

time /series

更新式 $\underbrace{\mathbf{w}_t = \mathbf{w}_{t-1}}_{} + \dfrac{\lambda^{-1}}{1 + \lambda^{-1}\mathbf{u}_t^{\mathrm{T}}\mathbf{P}_{t-1}\mathbf{u}_t} e_t\mathbf{P}_{t-1}\mathbf{u}_t$

再帰的に t-2, t-3 と展開 → 過去の $\mathbf{u}_{t-1}, \mathbf{u}_{t-2}, \ldots$ で構成

<u>それまでの入力情報が最適な $\mathbf{w}_{t-1}$ に集約されている</u>

他のトピック

非線形
- Volterra filter: 弱い非線形性に良い
- カーネルベースの適応フィルタ
  - 変換された入力ベクトル + 線形モデル

ロバスト性: 外乱（雑音）への対応
- Huber の M 推定: コスト関数の変更
- 雑音検出器: フィルタ更新を一時停止

高速化: サブバンド処理
- 複素表現パラメータ

ブラインド推定: 教師なし学習

実習・実装

LMS・NLMS 辺りは簡単
- Octave ならかなり手軽に試せる
- MIT HRTF などのインパルス応答を用いてシミュレート可能
  - https://sound.media.mit.edu/resources/KEMAR.html
- 雑音が重畳する場合など，パラメータを変更して色々試してみましょう

エコーキャンセラの設定: 参照信号をキャンセルするために伝達経路（フィルタ）を推定しているのであって，「（よくわからん）雑音をとにかく抑圧したい」というわけでない

参考文献

- 浅野太「音のアレイ信号処理」
- S. ヘイキン（武部 幹訳）「適応フィルタ入門」
- S. Haykin, “Adaptive Filter Theory fifth edition”

- 日本音響学会「音のなんでも小辞典」
- Albert S. Bregman: “Auditory Scene Analysis The Perceptual Organization of Sound”