逐次最小二乗法
Recursive Least Squares

大阪府立大学 工学研究科
清水　悠生

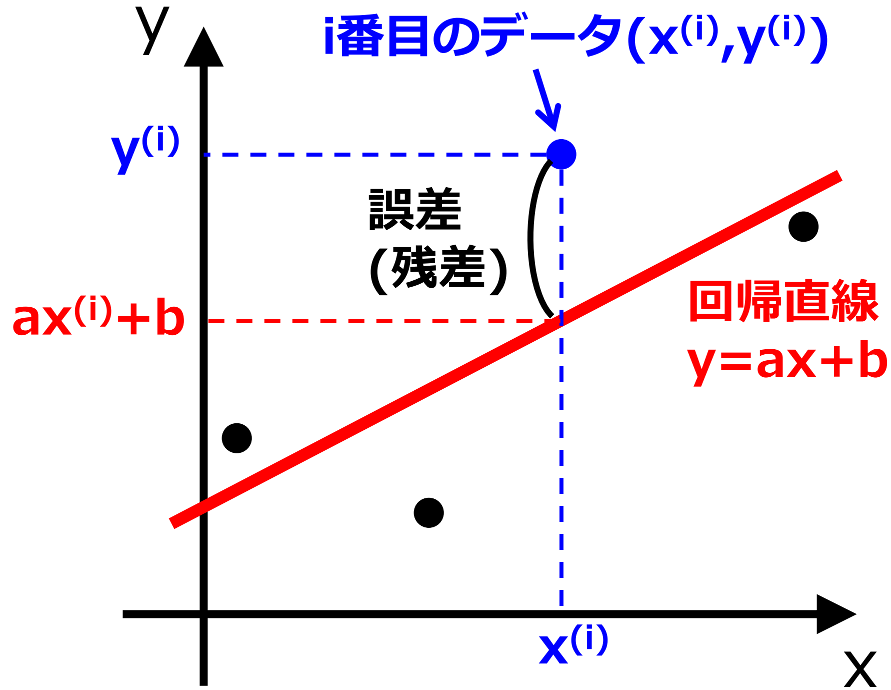
(復習)最小二乗法で扱う誤差関数

- ✓ 誤差関数を誤差の2乗の和とし，誤差関数が最小となるような係数a, bを計算する方法が最小二乗法

最小二乗法で扱う誤差関数E(a,b)

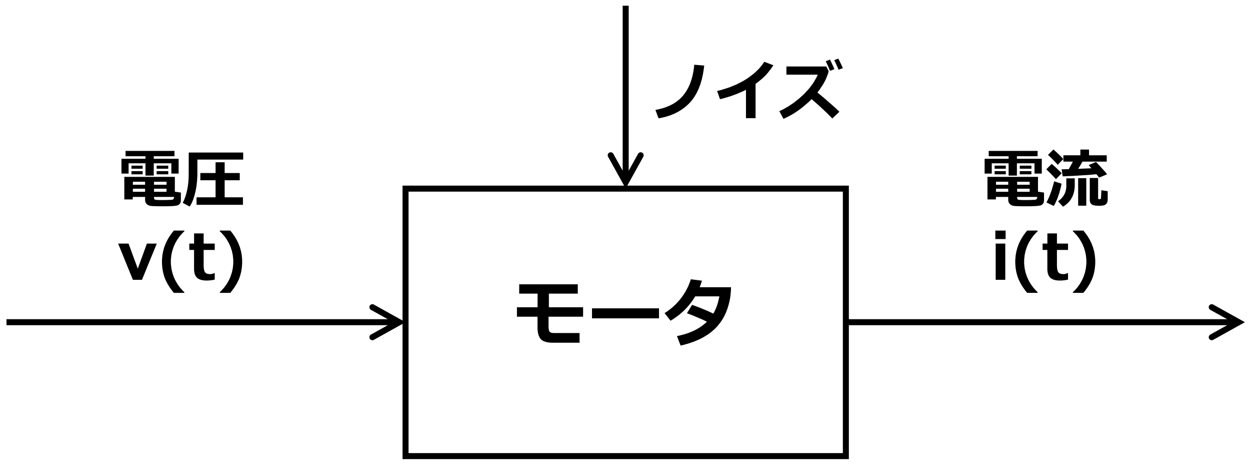
$$\underbrace{E(a,b)}_{\text{aとbの2変数関数}} = \sum_{i=1}^{n} \frac{1}{2} \underbrace{\left(y^{(i)} - \left(ax^{(i)} + b\right)\right)^2}_{\text{誤差の2乗}}$$

n個のデータを仮定



時系列データの場合は？

- ✓ 例えば，ある時刻 t において
  モータに電圧 v(t) をかけたとき電流 i(t) が流れたとする
- ✓ **ある一定時間(e.g. t=0,⋯,n)の電圧と電流の関係から**
  モータの抵抗値やインダクタンスなどを同定したい！
  ⇒**逐次最小二乗法**が有効！

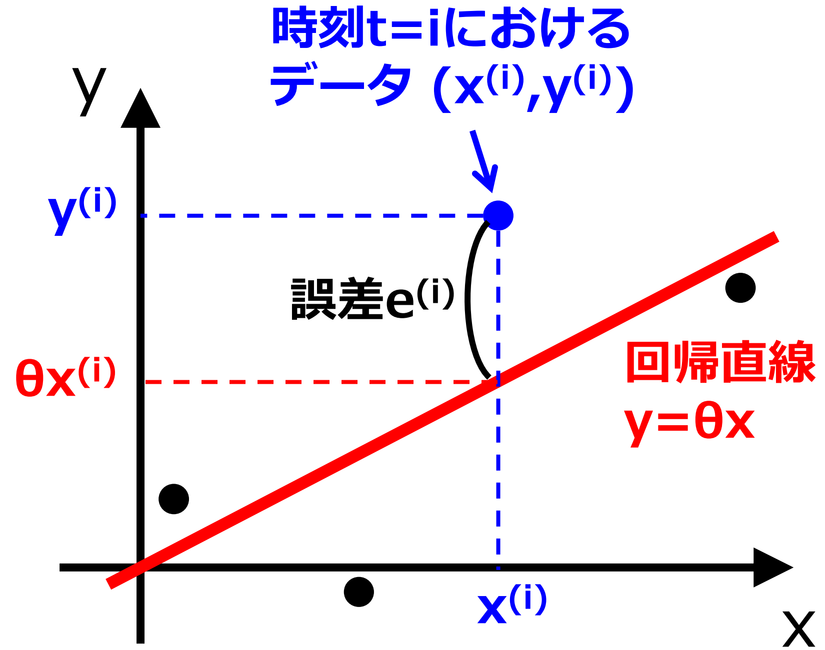


時系列データの場合の問題設定

- ✓ 時刻 t=i (i=0,1,⋯)における入力出力データ($x^{(i)}$,$y^{(i)}$)が得られた場合を仮定する
- ✓ xとyの関係は比例係数 θ の比例関係であると仮定（ただし出力データyにはノイズがのっている）
- ✓ この θ を最小二乗法で同定する！

時刻t=iにおける回帰誤差$e^{(i)}$

$$e^{(i)} = y^{(i)} - \theta x^{(i)}$$



時系列データの場合の2つ問題点

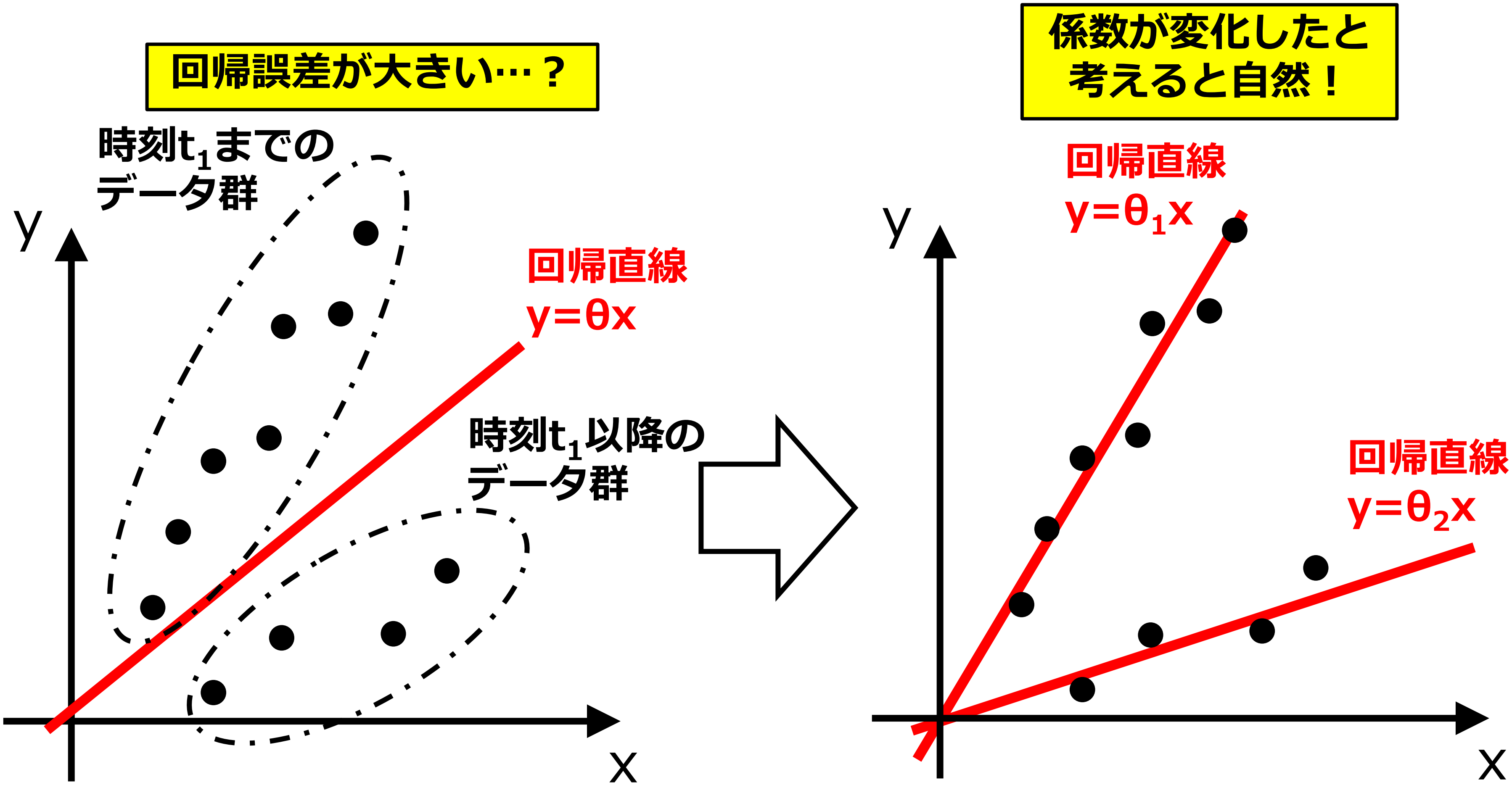
✓ 時系列データを扱う際には下記の2つの問題に注意

① 求めたい回帰係数（パラメータ）の時刻による変動

② メモリ（記憶領域）の限界

問題①：パラメータの変動

- ✓ **求めたい回帰係数（パラメータ）は変動する場合**がある
  ex)抵抗値の温度変化，インダクタンスの磁気飽和etc
- ✓ **パラメータの時間変動にすばやく追従する必要**がある！



過去のデータをどんどん忘れるように工夫！

✓ **忘却係数λを誤差関数に導入し,**
過去のデータの影響を小さくしていく⇒**逐次最小二乗法**

時刻t=iにおける回帰誤差e(i)

$$e^{(i)} = y^{(i)} - \theta x^{(i)}$$

通常の最小二乗法で扱う時刻t=nでの誤差関数E[n]

$$E[n] = \sum_{i=0}^{n} \frac{1}{2}\left(e^{(i)}\right)^2 = \frac{1}{2}\left(e^{(n)}\right)^2 + \frac{1}{2}\left(e^{(n-1)}\right)^2 + \cdots + \frac{1}{2}\left(e^{(0)}\right)^2$$

逐次最小二乗法で扱う時刻t=nでの誤差関数E[n]

$$E[n] = \sum_{i=0}^{n} \frac{1}{2}\lambda^{n-i}\left(e^{(i)}\right)^2 = \frac{1}{2}\underline{\lambda^0}\left(e^{(n)}\right)^2 + \frac{1}{2}\underline{\lambda^1}\left(e^{(n-1)}\right)^2 + \cdots + \frac{1}{2}\underline{\lambda^n}\left(e^{(0)}\right)^2$$

λ: 忘却係数 (0<λ<1)

過去のデータになればなるほど
忘却係数の次数が大きくなる
（値が小さくなる）

問題②：メモリの限界

- ✓ **時系列データはシステムが動作している間はずっと増え続ける**ため，その全てを記憶することは不可能

時刻t=nでの誤差関数E[n]

$$E[n]=\sum_{i=0}^{n}\frac{1}{2}\lambda^{n-i}\left(e^{(i)}\right)^2$$

nが大きくなればなるほど記憶しておくデータ数が膨大に！

- ✓ **直近のデータのみを使用してパラメータを計算できることが理想**

まとめると…

- ✓ 誤差関数E[n]から求まるパラメータθ[n]に関して漸化式を作ればよい！

$$E[n] = \sum_{i=0}^{n} \frac{1}{2} \lambda^{n-i} \left(e^{(i)}\right)^2$$ の最小化

⇒ $\theta[n]$が求められるが，
メモリの限界があるためこれは計算できない

⇒ **$\theta[n]$と$\theta[n-1]$の関係を数式化できれば（漸化式を作れば）直近のデータの計算だけでパラメータが求まる！**

- ✓ この漸化式を導出していきます

まずは普通に誤差関数を解いてみる

✓ **誤差関数の最小化**により
時刻 t=n における**パラメータの最適解 θ[n] を求める**

$$\min_{\theta} \quad E[n] = \sum_{i=0}^{n} \frac{1}{2} \lambda^{n-i} \left(e^{(i)}\right)^2 = \sum_{i=0}^{n} \frac{1}{2} \lambda^{n-i} \left(y^{(i)} - \theta x^{(i)}\right)^2$$

$$\Rightarrow \frac{\partial}{\partial \theta} E[n] = \sum_{i=0}^{n} \left(y^{(i)} - \theta[n] x^{(i)}\right) \lambda^{n-i} \left(-x^{(i)}\right)$$

微分して0

$$= (\boldsymbol{Y}[n] - \theta[n] \boldsymbol{X}[n])^{\mathrm{T}} \boldsymbol{\Lambda}[n] (-\boldsymbol{X}[n])$$

行列で表現

$$= -\boldsymbol{Y}[n]^{\mathrm{T}} \boldsymbol{\Lambda}[n] \boldsymbol{X}[n] + \theta[n] \boldsymbol{X}[n]^{\mathrm{T}} \boldsymbol{\Lambda}[n] \boldsymbol{X}[n] = 0$$

展開

$$\Leftrightarrow \boxed{\theta[n] = \left(\boldsymbol{X}[n]^{\mathrm{T}} \boldsymbol{\Lambda}[n] \boldsymbol{X}[n]\right)^{-1} \boldsymbol{Y}[n]^{\mathrm{T}} \boldsymbol{\Lambda}[n] \boldsymbol{X}[n]}$$

逆行列を左から
かける(正則と仮定)

ただし，時刻 t=n における各ベクトル・行列は次式のように定義

$$\boldsymbol{Y}[n] = \begin{bmatrix} y^{(0)} \\ \vdots \\ y^{(n)} \end{bmatrix}, \boldsymbol{X}[n] = \begin{bmatrix} x^{(0)} \\ \vdots \\ x^{(n)} \end{bmatrix}, \boldsymbol{\Lambda}[n] = \begin{bmatrix} \lambda^0 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda^n \end{bmatrix}$$
(n×n行列)

目標の再確認

✓ 時刻 t=n におけるパラメータの最適解 θ[n] と
時刻 t=n-1 におけるパラメータの最適解 θ[n-1] との
関係を表す漸化式を導出することが最終目標

$$\theta[n] = \left(\boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n]\right)^{-1}\boldsymbol{Y}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n]$$

$\Updownarrow$ **両者の関係を数式化する！**

$$\theta[n-1] = \left(\boldsymbol{X}[n-1]^{\mathrm{T}}\boldsymbol{\Lambda}[n-1]\boldsymbol{X}[n-1]\right)^{-1}\boldsymbol{Y}[n-1]^{\mathrm{T}}\boldsymbol{\Lambda}[n-1]\boldsymbol{X}[n-1]$$

漸化式の導出(1/3)

✓ θ[n]を変形する

$$\theta[n] = \underbrace{\left(\boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n]\right)^{-1}}_{=\, p[n]}\boldsymbol{Y}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n]$$

$= p[n]$**(スカラ)とおく**

$$\begin{aligned}
p[n] &= \left(\boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n]\right)^{-1} \\
&= \left(\sum_{i=0}^{n}\lambda^{n-i}\left(x^{(i)}\right)^2\right)^{-1} \\
&= \left(\sum_{i=0}^{n-1}\lambda^{n-i}\left(x^{(i)}\right)^2 + \lambda^0\left(x^{(n)}\right)^2\right)^{-1} \\
&= \left(\lambda\sum_{i=0}^{n-1}\lambda^{(n-1)-i}\left(x^{(i)}\right)^2 + \left(x^{(n)}\right)^2\right)^{-1} \\
&= \left(\lambda\left(p[n-1]\right)^{-1} + \left(x^{(n)}\right)^2\right)^{-1} = \underline{\frac{p[n-1]}{\lambda + p[n-1]\left(x^{(n)}\right)^2}}
\end{aligned}$$

要素の和に変形

t=nの要素を外に出す

第1項目をn-1で表現

漸化式の導出(2/3)

✓ つづき

$$\theta[n] = \left(\boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n]\right)^{-1}\underline{\boldsymbol{Y}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n]}$$

$= q[n]$**(スカラ)とおく**

$$q[n] = \boldsymbol{Y}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n]$$

$$= \sum_{i=0}^{n} y^{(i)}\lambda^{n-i}x^{(i)}$$

要素の和に変形

$$= \sum_{i=0}^{n-1} y^{(i)}\lambda^{n-i}x^{(i)} + y^{(n)}\lambda^{0}x^{(n)}$$

t=nの要素を外に出す

$$= \lambda\sum_{i=0}^{n-1} y^{(i)}\lambda^{(n-1)-i}x^{(i)} + y^{(n)}x^{(n)}$$

第1項目をn-1で表現

$$= \underline{\lambda q[n-1] + y^{(n)}x^{(n)}}$$

漸化式の導出(3/3)

✓ ついに漸化式が求まる！

$$\begin{aligned}\theta[n] &= \left(\boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n]\right)^{-1}\boldsymbol{Y}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n] \\ &= \underline{p[n]}\,\underline{q[n]} \\ &= \underline{\frac{p[n-1]}{\lambda + p[n-1](x^{(n)})^2}}\,\underline{\left(\lambda q[n-1] + y^{(n)}x^{(n)}\right)} \\ &= \frac{\lambda p[n-1]q[n-1] + p[n-1]y^{(n)}x^{(n)}}{\lambda + p[n-1](x^{(n)})^2} \\ &= \frac{\lambda\theta[n-1] + p[n-1]y^{(n)}x^{(n)}}{\lambda + p[n-1](x^{(n)})^2} \\ &= \theta[n-1] + \frac{p[n-1]x^{(n)}}{\lambda + p[n-1](x^{(n)})^2}\left(y^{(n)} - \theta[n-1]x^{(n)}\right)\end{aligned}$$

前ページの結果を代入

展開

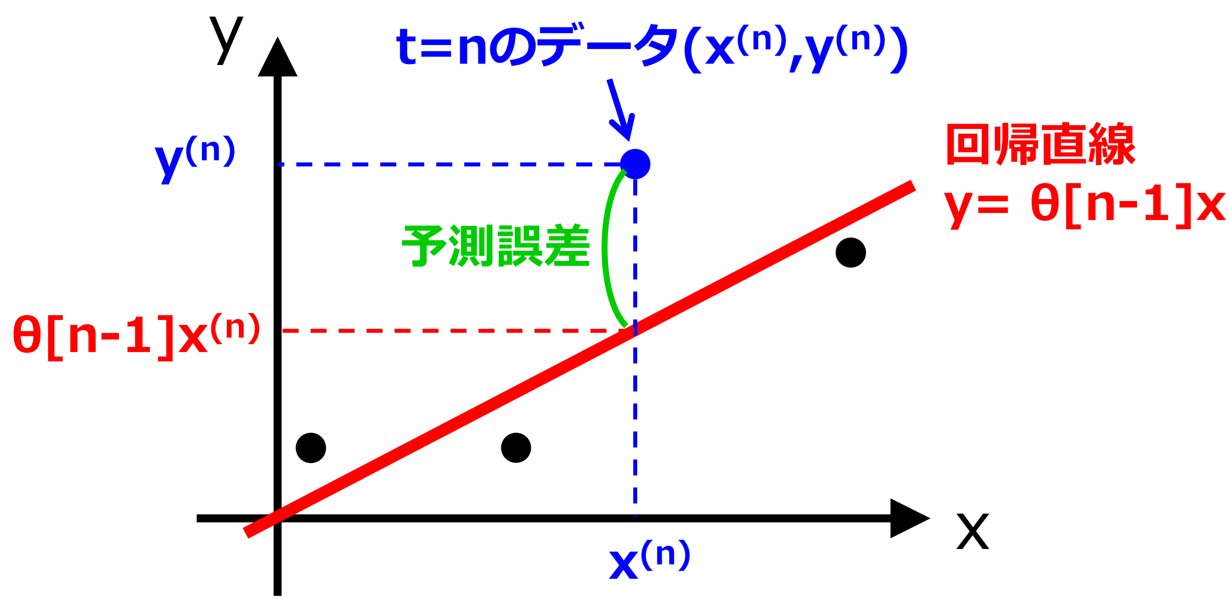
θ[n-1]を作る

一つ前のパラメータθ[n-1]とt=nのデータだけでθ[n]が導出可能に！
ただし，p[n-1]も漸化式から計算しておく必要あり

導出した漸化式の解釈

- ✓ 完成した漸化式はフィードバック制御のように解釈可能

$$\theta[n] = \underbrace{\theta[n-1]}_{\text{1つ前のパラメータ}} + \underbrace{\frac{p[n-1]x^{(n)}}{\lambda + p[n-1](x^{(n)})^2}}_{\text{λで決まるゲイン(誤差をどれだけ反映するか)}} \underbrace{\left(y^{(n)} - \theta[n-1]x^{(n)}\right)}_{\text{1つ前のパラメータによる予測誤差}}$$

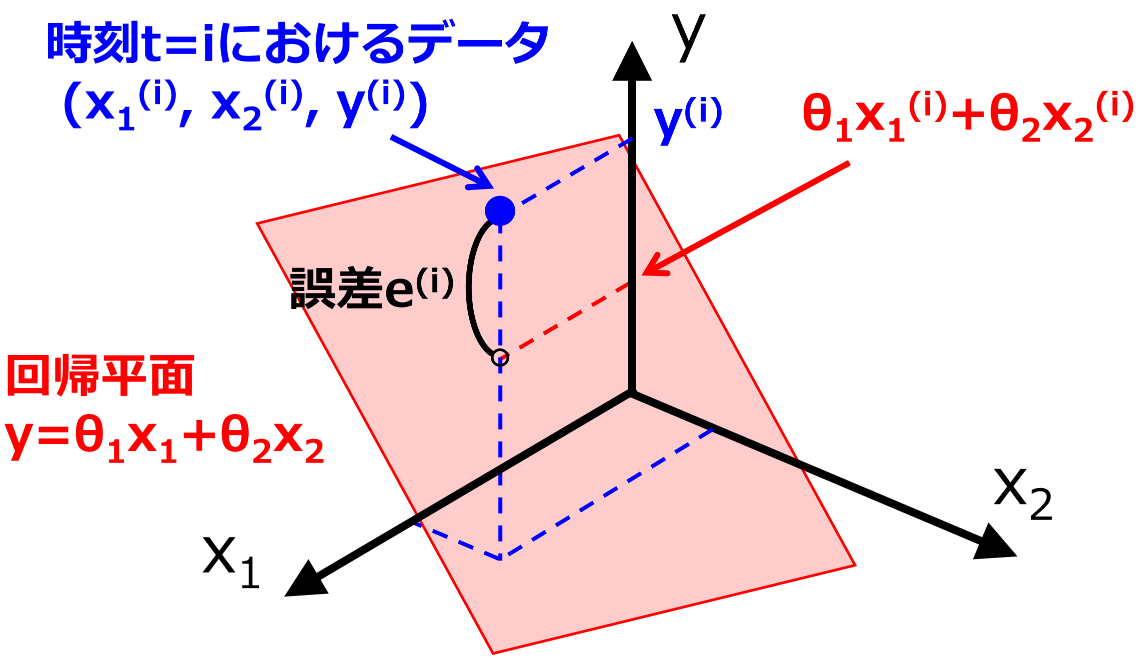


単回帰から重回帰への拡張

- ✓ **複数パラメータ θ_1, θ_2 を同時に同定する場合**を考える
- ✓ 回帰誤差は下記の通り

時刻t=iにおける回帰誤差$e^{(i)}$

$$e^{(i)} = y^{(i)} - \theta_1 x_1^{(i)} - \theta_2 x_2^{(i)}$$



まずは誤差関数の最小化

- ✓ パラメータ最適解 $\theta_1[n],\theta_2[n]$ を求めると
  単回帰の時と同じ解が得られる

$$\min_{\boldsymbol{\Theta}} \quad E[n] = \frac{1}{2}\left(\boldsymbol{Y}[n] - \boldsymbol{X}[n]\boldsymbol{\Theta}[n]\right)^{\mathrm{T}}\boldsymbol{\Lambda}[n]\left(\boldsymbol{Y}[n] - \boldsymbol{X}[n]\boldsymbol{\Theta}[n]\right)$$

$$\Rightarrow \frac{\partial}{\partial \boldsymbol{\Theta}} E[n] = \left(-\boldsymbol{X}[n]\right)^{\mathrm{T}}\boldsymbol{\Lambda}[n]\left(\boldsymbol{Y}[n] - \boldsymbol{X}[n]\boldsymbol{\Theta}[n]\right)$$

微分して0

$$= -\boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{Y}[n] + \boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n]\boldsymbol{\Theta}[n] = \boldsymbol{0}$$

展開

$$\Leftrightarrow \boxed{\boldsymbol{\Theta}[n] = \left(\boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n]\right)^{-1}\boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{Y}[n]}$$

逆行列を左からかける(正則と仮定)

ただし，時刻 t=n における各行列は次式のように定義

$$\boldsymbol{Y}[n] = \begin{bmatrix} y^{(0)} \\ \vdots \\ y^{(n)} \end{bmatrix} \textbf{(n×1)}, \boldsymbol{X}[n] = \begin{bmatrix} x_1^{(0)} & x_2^{(0)} \\ \vdots & \vdots \\ x_1^{(n)} & x_2^{(n)} \end{bmatrix} \textbf{(n×2)}$$

$$\boldsymbol{\Theta}[n] = \begin{bmatrix} \theta_1[n] \\ \theta_2[n] \end{bmatrix} \textbf{(2×1)}, \boldsymbol{\Lambda}[n] = \begin{bmatrix} \lambda^0 & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \lambda^n \end{bmatrix} \textbf{(n×n)}$$

漸化式の導出(1/4)

✓ Θ[n]を変形する

$$\boldsymbol{\Theta}[n] = \underline{\left(\boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n]\right)^{-1}}\boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{Y}[n]$$

$= \boldsymbol{P}[n]$ **(2×2)とおく**

$$\boldsymbol{P}[n] = \left(\boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n]\right)^{-1}$$

$$= \left(\sum_{i=0}^{n} \lambda^{n-i}\boldsymbol{x}^{(i)\mathrm{T}}\boldsymbol{x}^{(i)}\right)^{-1} \quad \left(\boldsymbol{x}^{(i)} = \begin{bmatrix} x_1^{(i)} & x_2^{(i)} \end{bmatrix}\right)$$

要素の和に変形

$$= \left(\lambda\sum_{i=0}^{n-1} \lambda^{(n-1)-i}\boldsymbol{x}^{(i)\mathrm{T}}\boldsymbol{x}^{(i)} + \lambda^0\boldsymbol{x}^{(n)\mathrm{T}}\boldsymbol{x}^{(n)}\right)^{-1}$$

t=nの要素を外に出す

$$= \left(\lambda(\boldsymbol{P}[n-1])^{-1} + \boldsymbol{x}^{(n)\mathrm{T}}\boldsymbol{x}^{(n)}\right)^{-1}$$

第1項目をn-1で表現

漸化式の導出(2/4)

✓ P[n]のつづき

下記の公式を用いる（ひたすら要素計算すると証明できます）

$$\left(\boldsymbol{A}^{-1}+\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}\right)^{-1}=\boldsymbol{A}-\frac{\boldsymbol{A}\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}}{1+\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{B}}$$

$\boldsymbol{A}$：M×Mの正定値行列
$\boldsymbol{B}$：M×1行列

公式に代入して

$$\boldsymbol{P}[n]=\left(\lambda(\boldsymbol{P}[n-1])^{-1}+\boldsymbol{x}^{(n)^{\mathrm{T}}}\boldsymbol{x}^{(n)}\right)^{-1}$$

$$=\frac{1}{\lambda}\boldsymbol{P}[n-1]-\frac{\frac{1}{\lambda}\boldsymbol{P}[n-1]\boldsymbol{x}^{(n)^{\mathrm{T}}}\boldsymbol{x}^{(n)}\frac{1}{\lambda}\boldsymbol{P}[n-1]}{1+\boldsymbol{x}^{(n)}\frac{1}{\lambda}\boldsymbol{P}[n-1]\boldsymbol{x}^{(n)^{\mathrm{T}}}}$$

公式を適用

$$=\frac{1}{\lambda}\left(\boldsymbol{P}[n-1]-\frac{\boldsymbol{P}[n-1]\boldsymbol{x}^{(n)^{\mathrm{T}}}\boldsymbol{x}^{(n)}\boldsymbol{P}[n-1]}{\lambda+\boldsymbol{x}^{(n)}\boldsymbol{P}[n-1]\boldsymbol{x}^{(n)^{\mathrm{T}}}}\right)$$

1/λでくくる

$$=\frac{\boldsymbol{P}[n-1]}{\lambda+\boldsymbol{x}^{(n)}\boldsymbol{P}[n-1]\boldsymbol{x}^{(n)^{\mathrm{T}}}}$$

一つにまとめる

最後の変形では下記の関係を用いた（これもひたすら計算です）

$$\boldsymbol{P}[n-1]\boldsymbol{x}^{(n)}\boldsymbol{P}[n-1]\boldsymbol{x}^{(n)^{\mathrm{T}}}=\boldsymbol{P}[n-1]\boldsymbol{x}^{(n)^{\mathrm{T}}}\boldsymbol{x}^{(n)}\boldsymbol{P}[n-1]$$

漸化式の導出(3/4)

✓ つづき

$$\boldsymbol{\Theta}[n] = \left(\boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n]\right)^{-1}\underline{\boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{Y}[n]}$$

$= \boldsymbol{Q}[n]$(2×1行列)とおく

$$\boldsymbol{Q}[n] = \boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{Y}[n]$$

$$= \sum_{i=0}^{n} \lambda^{n-i} \boldsymbol{x}^{(i)\mathrm{T}} y^{(i)}$$

要素の和に変形

$$= \lambda \sum_{i=0}^{n-1} \lambda^{(n-1)-i} \boldsymbol{x}^{(i)\mathrm{T}} y^{(i)} + \lambda^{0} \boldsymbol{x}^{(n)\mathrm{T}} y^{(n)}$$

t=nの要素を外に出す

$$= \underline{\lambda \boldsymbol{Q}[n-1] + \boldsymbol{x}^{(n)\mathrm{T}} y^{(n)}}$$

第1項目をn-1で表現

漸化式の導出(4/4)

✓ 単回帰と同じ形式の漸化式が導出できる

$$\boldsymbol{\Theta}[n] = \left(\boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{X}[n]\right)^{-1}\boldsymbol{X}[n]^{\mathrm{T}}\boldsymbol{\Lambda}[n]\boldsymbol{Y}[n]$$

$$= \boldsymbol{P}[n]\boldsymbol{Q}[n]$$

$$= \frac{\boldsymbol{P}[n-1]}{\lambda + \boldsymbol{x}^{(n)}\boldsymbol{P}[n-1]\boldsymbol{x}^{(n)\mathrm{T}}}\left(\lambda\boldsymbol{Q}[n-1] + \boldsymbol{x}^{(n)\mathrm{T}}y^{(n)}\right)$$

前ページの結果を代入

$$= \frac{\lambda\boldsymbol{P}[n-1]\boldsymbol{Q}[n-1] + \boldsymbol{P}[n-1]\boldsymbol{x}^{(n)\mathrm{T}}y^{(n)}}{\lambda + \boldsymbol{x}^{(n)}\boldsymbol{P}[n-1]\boldsymbol{x}^{(n)\mathrm{T}}}$$

展開

$$= \frac{\lambda\boldsymbol{\Theta}[n-1] + \boldsymbol{P}[n-1]\boldsymbol{x}^{(n)\mathrm{T}}y^{(n)}}{\lambda + \boldsymbol{x}^{(n)}\boldsymbol{P}[n-1]\boldsymbol{x}^{(n)\mathrm{T}}}$$

Θ[n-1]を作る

$$= \underbrace{\boldsymbol{\Theta}[n-1]}_{\text{1つ前のパラメータ}} + \underbrace{\frac{\boldsymbol{P}[n-1]\boldsymbol{x}^{(n)\mathrm{T}}}{\lambda + \boldsymbol{x}^{(n)}\boldsymbol{P}[n-1]\boldsymbol{x}^{(n)\mathrm{T}}}}_{\lambda\text{で決まるゲイン}}\underbrace{\left(y^{(n)} - \boldsymbol{x}^{(n)}\boldsymbol{\Theta}[n-1]\right)}_{\text{1つ前のパラメータによる予測誤差}}$$

1つ前のパラメータ

λで決まるゲイン

1つ前のパラメータによる予測誤差

逐次最小二乗法の活用例

✓ 電圧方程式からIPMSMのインダクタンス等のパラメータをオンラインで同定することが可能

定常状態の電圧方程式(d軸)

$$\underbrace{v_d}_{Y} = R_a i_d + -\omega L_q i_q = \underbrace{\begin{bmatrix} R_a & L_q \end{bmatrix}}_{\text{パラメータ}\Theta} \underbrace{\begin{bmatrix} i_d \\ -\omega i_q \end{bmatrix}}_{X}$$

[1] Morimoto et al., “Mechanical Sensorless Drives of IPMSM With Online Parameter Identification,” IEEE Trans. Ind. Appl., vol. 42, no. 5, pp. 1241-1248, 2006