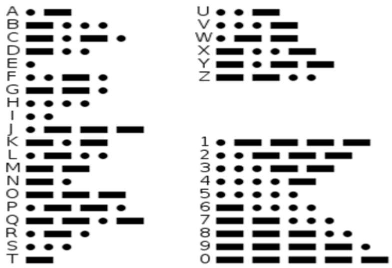
マルチメディア信号処理
前半のまとめ

モールス符号



Message ← Morse Code

SOS ← ・・・ーーー・・・

最も簡単な圧縮は、モールス信号

文字	モールス符号 i	発生確率 P_i	長さ L_i
A	・ー	0.064	6
B	ー・・・	0.013	10
C	ー・ー・	0.022	12
Space	←→ 符号長	0.186	1

【二値化】それぞれの文字を「・」か「ー」で置換する

【 圧 縮 】通報全体の符号長を出来るだけ短くしたい

頻出する文字には短い符号を，
さもなければ長い符号を割り当てる．

圧縮の基本 ハフマン符号化

通報＝｛ ABC_AAB_AABA ｝ の場合

文字	自然2進数	ﾊﾌﾏﾝ符号	発生頻度
A	00 [2 bit]	1 [1 bit] （短い）	6 （頻出）
B	01 [2 bit]	00 [2 bit] ↑	3 ↑
_	11 [2 bit]	010 [3 bit] ↓	2 ↓
C	10 [2 bit]	011 [3 bit] （長い）	1 （希少）

それぞれの文字を0と1の符号で置き換える（二値化する）

圧縮後は…

通報＝｛ ABC_AAB_AABA ｝ の場合

文字	ﾊﾌﾏﾝ符号
A	1 [1 bit]
B	00 [2 bit]
_	010 [3 bit]
C	011 [3 bit]

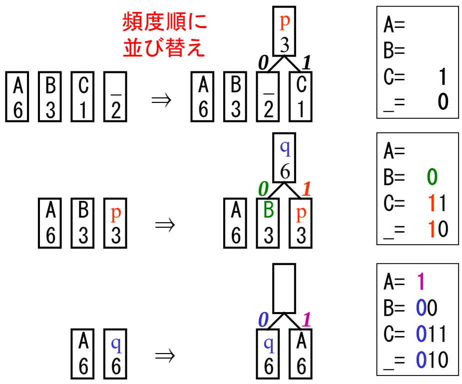
→ 1BC_AAB_AABA
→ 100C_AAB_AABA
→ 10001101・・・AB_AABA
→ 100011010110001011001
【21 bit 圧縮後】

圧縮

→ 000110110000011100000100
【24 bit 圧縮前】

→ ”符号長”を各自計算すること

ハフマン符号を 作ってみよう



各文字の発生頻度が分かれば、圧縮できる。

課題

【問題１】一番圧縮できる通報を選べ.

- 通報①　{ABC_ABC_ABC_}
- 通報②　{AABBBC_AC_C_}
- 通報③　{ABA_ABABCACA}

【問題2】どのような場合に効率良く圧縮できるか？

【問題3】通報①は，全く圧縮出来ないか？

自己情報量 とは何か？

発生確率	自己情報量	ハフマン符号	
$P_{(1)}$= 5/11	$-\log_2 P_{(1)}$= 1.14 bit	1	1 bit
$P_{(2)}$= 3/11	$-\log_2 P_{(2)}$= 1.87 bit	00	2 bit
$P_{(3)}$= 2/11	$-\log_2 P_{(3)}$= 2.46 bit	011	3 bit
$P_{(4)}$= 1/11	$-\log_2 P_{(4)}$= 3.46 bit	010	3 bit

発生確率 ⇒ **自己情報量** ⇒ 符号長の推定

エントロピー とは何か？

発生確率	重み×自己情報量	ハフマン符号	
$P_{(1)}$= 5/11	$-P_{(1)}\log_2 P_{(1)}$= 0.52	1	1bit×5/11= 0.45
$P_{(2)}$= 3/11	$-P_{(2)}\log_2 P_{(2)}$= 0.51	00	2bit×3/11= 0.55
$P_{(3)}$= 2/11	$-P_{(3)}\log_2 P_{(3)}$= 0.45	011	3bit×2/11= 0.55
$P_{(4)}$= 1/11	$-P_{(4)}\log_2 P_{(4)}$= 0.31	010	3bit×1/11= 0.27
	合計 1.79 bit/文字 **エントロピ**（平均情報量）	合計 1.82 bit/文字 **平均符号長**	

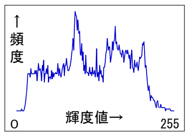
自己情報量の重みつき和 ⇒ エントロピ ⇒ 平均符号長の推定

ハフマン符号化だけでは不十分

64 KB （100%）
原画像 x(n)
↓
ハフマン符号化
↓
圧縮データ
60 KB （94%）

→ m
↓ n

原画像
x(n,m)



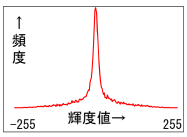
ヒストグラム

微分を併用して効果的に圧縮

64 KB （100%）
原画像 x(n)
↓
微分処理
y(n)= x(n) − x(n−1)
↓
ハフマン符号化
↓
圧縮データ
54 KB （84%）＜ 94%

→ m
↓ n

予測誤差+128
y(n,m) = x(n,m) − x(n,m−1)



ヒストグラム

予測の実例　微分と積分で元通り！

原信号

$$[x(0)\ \ x(1)\ \ x(2)] = [10\ \ 12\ \ 11]$$

予測処理（微分）

$$y(0) = x(0) - x(-1) = 10 - 0 = 10$$
$$y(1) = x(1) - x(0) = 12 - 10 = 2$$
$$y(2) = x(2) - x(1) = 11 - 12 = -1$$

予測誤差

再生方法（積分）

$$\underline{x}(0) = y(0) + \underline{x}(-1) = 10 + 0 = 10$$
$$\underline{x}(1) = y(1) + \underline{x}(0) = 2 + 10 = 12$$
$$\underline{x}(2) = y(2) + \underline{x}(1) = -1 + 12 = 11$$

課題

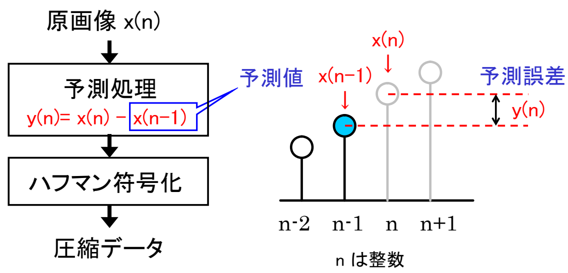
1．以下の２つのうち、圧縮できる方を選べ。

[x(0)　x(1)　…　x(5)] = [2　2　2　1　1　1]
[x(0)　x(1)　…　x(5)] = [1　0　1　0　1　0]

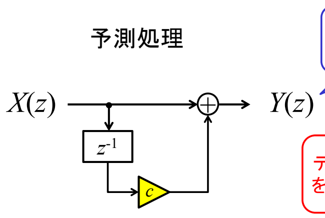
2．予測処理　y(n) = x(n) − x(n−1)　を施せ。

3．予測の結果を、ハフマン符号化せよ。

微分 は 予測　（0次の外挿）



最適な予測器 を設計しよう



データ量を最小化

$H \propto \log_2 \sigma_y^2$ （エントロピ，分散）

$$y(n) = x(n) - c \cdot x(n-1)$$

c を求める

$$\sigma_y^2 = \frac{1}{N}\sum_n \{y(n) - m_y\}^2$$

$$m_y = \frac{1}{N}\sum_n y(n)$$

最適な予測器 を設計しよう

以下の予測処理，

y(n) = x(n) + ***c*** x(n−1)

において，与えられた入力信号 “x(n)”

に対する最適な “***c***” の値を求めたい．

予測後のデータ量 ∝ log(予測誤差の分散)

⇒ 予測誤差の分散を最小化する

最適化の方法

1．予測誤差 “y(n)” の分散値を “*c*” で表す．　$m_y = \frac{1}{N}\sum_n y(n) = 0$

$$\sigma_y^2 = \frac{1}{N}\sum_{\forall n} \{y(n) - m_y\}^2,\quad \text{以下、平均 } m_y = 0 \text{ のとき}$$

$$= \frac{1}{N}\sum_{\forall n} \{x(n) + c \cdot x(n-1)\}^2$$

2．得られた分散値を最小とする “*c*” は次式を満たす．

$$\frac{\partial \sigma_y^2}{\partial c} = 0$$

以上を解けば最適係数値が決定される．

3．計算すると

$$\frac{\partial}{\partial c}\sigma_y^2 = \frac{1}{N}\sum_{\forall n} 2\{x(n) + c \cdot x(n-1)\} \cdot \frac{\partial}{\partial c}\{x(n) + c \cdot x(n-1)\} = 0$$

$$\sum_{\forall n} \{x(n) + c \cdot x(n-1)\} \cdot \frac{\partial}{\partial c}\{x(n) + c \cdot x(n-1)\} = 0$$

$$\sum_{\forall n} \{x(n) + c \cdot x(n-1)\} \cdot x(n-1) = 0$$

$$\sum_{\forall n} x(n)x(n-1) + c \cdot \sum_{\forall n} x^2(n-1) = 0$$

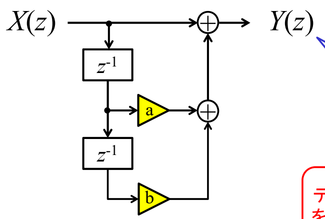
4．以上より、

$$\boxed{c = \frac{-R_1}{R_0}}$$

最適な予測係数は
自己相関で決まる

$$\begin{cases} R_0 = \sum_n x^2(n-1) \\ R_1 = \sum_n x(n)x(n-1) \end{cases}$$

最適な予測器を設計しよう（K= 2次）



$$y(n) = x(n) + a \cdot x(n-1) + b \cdot x(n-2)$$

a, b を求める

分散が最小となるよう a ,b を最適化する

⇩

データ量を最小化 $H \propto \log_2 \sigma_y^2$

エントロピ　分散

$$\sigma_y^2 = \frac{1}{N}\sum_n \{y(n) - m_y\}^2$$

$$m_y = \frac{1}{N}\sum_n y(n)$$

最適な予測器（パラメータ 2つ）

以下の予測処理,

y(n) = x(n) +***a*** x(n−1) +***b*** x(n−2)

において，与えられた入力信号 “x(n)”

に対する最適な “***a, b***” の値を求めたい.

⇩

予測後のデータ量 ∝ log(予測誤差の分散)

⇒ 予測誤差の分散を最小化する

最適化の方法（パラメータ2つ）

1．予測誤差 “y(n)” の分散値を “*a, b*” で表す. $m_y = \frac{1}{N}\sum_n y(n) = 0$

$$\sigma_y^2 = \frac{1}{N}\sum_n \{y(n) - m_y\}^2,\quad \text{以下、平均 } m_y = 0 \text{ のとき}$$

$$= \frac{1}{N}\sum_n \{x(n) + a \cdot x(n-1) + b \cdot x(n-2)\}^2$$

2．得られた分散値を最小とする “*a, b*” は次式を満たす.

$$\frac{\partial}{\partial a}\sigma_y^2 = 0,\quad \frac{\partial}{\partial b}\sigma_y^2 = 0$$

この方程式を解けば最適な係数が決まる

3．最初の方程式は、

$$f = x(n) + a \cdot x(n-1) + b \cdot x(n-2) \qquad (1)$$

とおくと、

$$\frac{\partial}{\partial a}\sigma_y^2 \; = \frac{\partial}{\partial a}\frac{1}{N}\sum_n f^2 \; = \frac{1}{N}\sum_n 2f \cdot \frac{\partial}{\partial a} f \; = 0$$

$$\therefore \sum_n f \cdot \frac{\partial}{\partial a} f \; = 0 \qquad (2)$$

$$(1) \to (2)$$

$$\sum_n \{x(n) + a \cdot x(n-1) + b \cdot x(n-2)\} \cdot x(n-1) \; = 0$$

$$\sum_n x(n)x(n-1) + a \cdot \sum_n x^2(n-1) + b \cdot \sum_n x(n-1)x(n-2) \; = 0$$

以上より、

$$\boxed{R_{01} + a \cdot R_{11} + b \cdot R_{12} \; = 0}$$

但し,

$$R_{ji} = \frac{1}{N}\sum_{n=0}^{N-1} x(n-j)x(n-i)$$

4．もう一つの方程式より、

$$\frac{\partial}{\partial b}\sigma_y^2 \; = \frac{1}{N}\sum_n 2f \cdot \frac{\partial}{\partial b} f \; = 0 \;\to\; \sum_n f \cdot \frac{\partial}{\partial b} f \; = 0$$

に

$$f = x(n) + a \cdot x(n-1) + b \cdot x(n-2)$$

を代入して、

$$\sum_n \{x(n) + a \cdot x(n-1) + b \cdot x(n-2)\} \cdot x(n-2) \; = 0$$

$$\sum_n x(n)x(n-2) + a \cdot \sum_n x(n-1)x(n-2) + b \cdot \sum_n x^2(n-2) \; = 0$$

以上より、

$$\boxed{R_{02} + a \cdot R_{12} + b \cdot R_{22} \; = 0}$$

但し,

$$R_{ji} = \frac{1}{N}\sum_{n=0}^{N-1} x(n-j)x(n-i)$$

5．得られた2つの方程式より

$$\begin{cases} R_{01} + a \cdot R_{11} + b \cdot R_{12} \; = 0 \\ R_{02} + a \cdot R_{12} + b \cdot R_{22} \; = 0 \end{cases}$$

自己相関が

$$\begin{cases} R_{11} = R_{22} = \Theta_0 \\ R_{01} = R_{12} = \Theta_1 \\ R_{02} = R_{13} = \Theta_2 \end{cases} \qquad \leftarrow \; R_{ji} = \frac{1}{N}\sum_{n=0}^{N-1} x(n-j)x(n-i)$$

を満たす場合、方程式は、

$$\begin{cases} \Theta_1 + a \cdot \Theta_0 + b \cdot \Theta_1 \; = 0 \\ \Theta_2 + a \cdot \Theta_1 + b \cdot \Theta_0 \; = 0 \end{cases} \;\Rightarrow\; \begin{pmatrix} \Theta_1 \\ \Theta_2 \end{pmatrix} + \begin{pmatrix} \Theta_0 & \Theta_1 \\ \Theta_1 & \Theta_0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

6. 以上より、

$$\begin{pmatrix} a \\ b \end{pmatrix} = -\begin{pmatrix} \Theta_0 & \Theta_1 \\ \Theta_1 & \Theta_0 \end{pmatrix}^{-1} \begin{pmatrix} \Theta_1 \\ \Theta_2 \end{pmatrix}$$

なので、

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{-1}{\Theta_0^2 - \Theta_1^2} \begin{pmatrix} \Theta_0 & -\Theta_1 \\ -\Theta_1 & \Theta_0 \end{pmatrix} \begin{pmatrix} \Theta_1 \\ \Theta_2 \end{pmatrix}$$

$$= \frac{-1}{\Theta_0^2 - \Theta_1^2} \begin{pmatrix} \Theta_0\Theta_1 - \Theta_1\Theta_2 \\ -\Theta_1\Theta_1 + \Theta_0\Theta_2 \end{pmatrix}$$

最適な予測係数は
自己相関で決まる

但し，

$$\Theta_k = \frac{1}{N}\sum_{n=0}^{N-1} x(n-j)x(n-i), \quad |j-i| = k$$

k次の自己相関

---

k次の相関係数を、

$$\Phi_k = \frac{\Theta_k}{\Theta_0}$$

但し，

$$\Theta_k = \frac{1}{N}\sum_{n=0}^{N-1} x(n-j)x(n-i), \quad |j-i| = k$$

と定義すると、

$$\begin{cases} a = \dfrac{\Theta_1\Theta_2 - \Theta_0\Theta_1}{\Theta_0^2 - \Theta_1^2} = \dfrac{-(1-\Phi_2)\Phi_1}{1-\Phi_1^2} \\ b = \dfrac{\Theta_1\Theta_1 - \Theta_0\Theta_2}{\Theta_0^2 - \Theta_1^2} = \dfrac{\Phi_1^2 - \Phi_2}{1-\Phi_1^2} \end{cases}$$

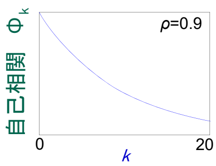
最適な予測係数は
自己相関係数
で決まる

AR(1)モデル に対する 最適予測

AR(1) model の自己相関は、

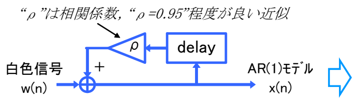
$$\Phi_k = \rho^k$$

と表される。このとき、



$$\begin{cases} a = \dfrac{-(1-\Phi_2)\Phi_1}{1-\Phi_1^2} = \dfrac{-(1-\rho^2)\rho}{1-\rho^2} = -\rho \\ b = \dfrac{\Phi_1^2 - \Phi_2}{1-\Phi_1^2} = \dfrac{\rho^2 - \rho^2}{1-\rho^2} = 0 \end{cases}$$

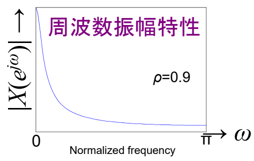
AR(1)モデルとは？

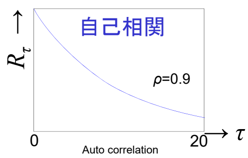


$$x(n) = w(n) + \rho \cdot x(n-1)$$

但し$w(n)$は平均値が零の白色信号

x(n) の
周波数振幅特性
と
自己相関関数
を求めよ





AR(1)モデルの周波数振幅特性と自己相関

$$X(e^{j\omega}) = \frac{1}{1-\rho\, e^{j\omega}}$$

$$|X(e^{j\omega})| = \frac{1}{\sqrt{(1-\rho\cdot\cos\omega)^2 + (-\sin\omega)^2}}$$

$$= \frac{1}{\sqrt{1+\rho^2 - 2\rho\cdot\cos\omega}}$$

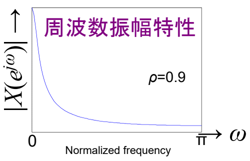
$$R_1 = \frac{1}{N}\sum_{n=0}^{N-1} x(n)x(n-1)$$

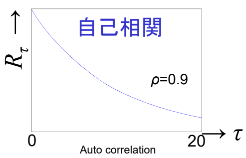
$$= \frac{1}{N}\sum_{n=0}^{N-1} \{\rho\cdot x(n-1) + w(n)\}x(n-1)$$

$$= \rho\cdot\frac{1}{N}\sum_{n=0}^{N-1} x^2(n-1) + \frac{1}{N}\sum_{n=0}^{N-1} w(n)x(n-1)$$

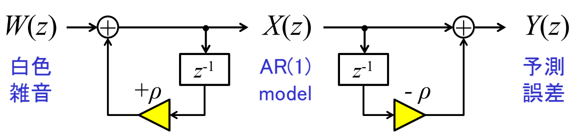
$$\approx \rho\cdot R_0$$

$$\therefore \boxed{R_k = \rho^k \cdot R_0}$$





AR(1) モデルの最適予測

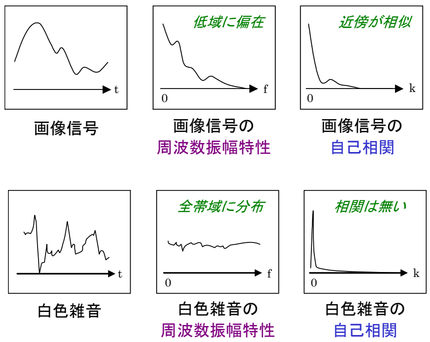


$$x(n) = w(n) + \rho\cdot x(n-1) \qquad y(n) = x(n) - \rho\cdot x(n-1)$$

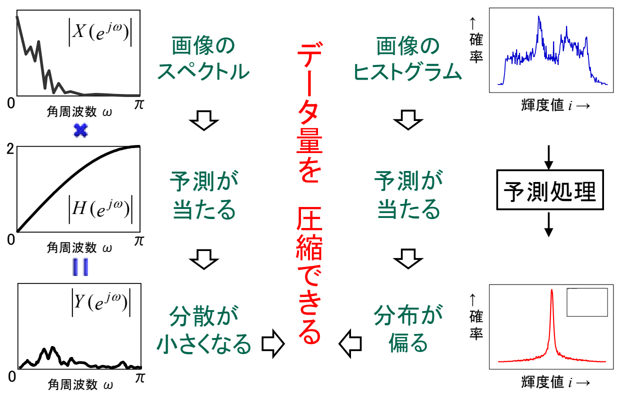
$$X(z) = \frac{1}{1-\rho\cdot z^{-1}}W(z) \qquad Y(z) = (1-\rho\cdot z^{-1})X(z)$$

$$Y(z) = W(z)$$

周波数振幅特性 と 自己相関



圧縮する ＝ 分散を小さくする



圧縮する ＝ 分散を小さくする

$X(e^{j\omega})$ 画像の スペクトル

データ量を 圧縮できる $H \propto \log_2 \sigma_y^2$

⇩ ⇧

$Y(e^{j\omega}) = H(e^{j\omega})X(e^{j\omega})$ 予測が 当たる

エントロピーと 分布の関係 $H_y = -\int_{-\infty}^{\infty} P(y)\log_2 P(y)dy$

⇩ ⇧

$\sigma_y^2 = \frac{1}{\pi}\int_{0^+}^{\pi} \left|Y(e^{j\omega})\right|^2 d\omega \rightarrow \min$ 分散が 小さくなる ⇨ 分布が 偏る $\sigma_y^2 = \int_{-\infty}^{\infty} P(y)\cdot y^2 dy$

分散 と エントロピ

確率密度関数 $P(x)$ が以下である信号 $x(n)$ について、
エントロピー H と 分散 σ^2 の関係を示せ。

$$P(x) = \begin{cases} 1/\Delta & for \quad |x| \le \Delta/2 \\ 0 & for \quad |x| > \Delta/2 \end{cases}$$

⇩

ヒント

$$H = -\int_{-\infty}^{\infty} P(x)\log_2 P(x)dx$$

$$\sigma^2 = \int_{-\infty}^{\infty} P(x)x^2 dx$$

分散を小さく
→ データ圧縮

⇨ $H \propto \log_2 \sigma^2$

分散 と エントロピ

$$\sigma^2 = \int_{-\infty}^{\infty} P(x)x^2 dx = 2\int_0^{\Delta/2} \frac{1}{\Delta}x^2 dx = \frac{2}{\Delta}\int_0^{\Delta/2} x^2 dx = \frac{2}{\Delta}\left[\frac{x^3}{3}\right]_0^{\Delta/2} = \frac{\Delta^2}{12}$$

$$H = -\int_{-\infty}^{\infty} P(x)\log_2 P(x)dx = -2\int_0^{\Delta/2} \frac{1}{\Delta}\log_2\frac{1}{\Delta}dx = \frac{2\log_2\Delta}{\Delta}\int_0^{\Delta/2} dx = \log_2\Delta$$

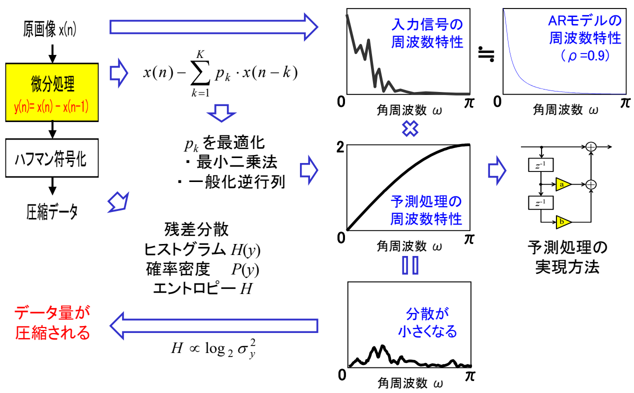
⇩

$$H = \log_2\Delta = \frac{1}{2}\left(\log_2\sigma^2 + \log_2 12\right)$$

分散を小さく
→ データ圧縮

⇨ $H \propto \log_2 \sigma^2$

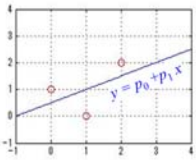
前半の 話の流れ



関連事項

線形回帰

課題　p_0, p_1 を求よ



			データ	直線上の点
n	x_n	y_n		$f(x)$
1	0	1		$p_0+p_1 \cdot 0$
2	1	0		$p_0+p_1 \cdot 1$
3	2	2		$p_0+p_1 \cdot 2$

←差→

方法１【最小二乗法】

$$I = \sum_n \{f(x_n) - y_n\}^2, \quad \frac{\partial I}{\partial p_k} = 0$$

方法２【一般化逆行列】

$$(\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T$$

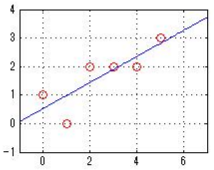
発展　２次の場合は？

$$y = p_0 + p_1 x + p_2 x^2$$

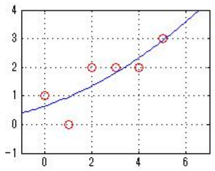
$\mathbf{P}$ =[1.5　-2.5　1.00]

次数を増やすと…

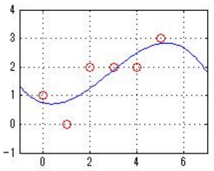
直線（１次）



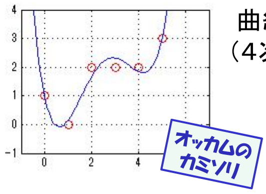
曲線（２次）



曲線（３次）



曲線（４次）



方法１【最小二乗法】

誤差を定義する

			データ	直線上の点
n	x_n	y_n		$f(x)$
1	0	1		$p_0+p_1 \cdot 0$
2	1	0		$p_0+p_1 \cdot 1$
3	2	2		$p_0+p_1 \cdot 2$

←差→

$$I = \{1 - (p_0+p_1 \cdot 0)\}^2 + \{0 - (p_0+p_1 \cdot 1)\}^2 + \{2 - (p_0+p_1 \cdot 2)\}^2$$

方程式を解く

$$\begin{cases} \dfrac{\partial I}{\partial p_0} = 0 \\ \dfrac{\partial I}{\partial p_1} = 0 \end{cases}$$

$$0 = 2\{1 - (p_0+p_1 \cdot 0)\}(-1) + 2\{0 - (p_0+p_1 \cdot 1)\}(-1) + 2\{2 - (p_0+p_1 \cdot 2)\}(-1)$$

$$0 = 2\{1 - (p_0+p_1 \cdot 0)\} \cdot (-0) + 2\{0 - (p_0+p_1 \cdot 1)\} \cdot (-1) + 2\{2 - (p_0+p_1 \cdot 2)\} \cdot (-2)$$

$$\begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

$$\therefore \begin{bmatrix} p_0 \\ p_1 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

方法１【最小二乗法】

1. 誤差を定義する

$$I = \sum_{n=1}^{N} \{f(x_n) - y_n\}^2 \quad \text{但し、} f(x) = \sum_{k=0}^{K} p_k x^k$$

2. 方程式を解く

$$\frac{\partial I}{\partial p_k} = 0, \quad k = 1,2,\cdots,K$$

$$\begin{cases} \dfrac{\partial I}{\partial p_0} = \sum_{n=1}^{N} 2 \cdot \{p_0 + p_1 \cdot x_n - y_n\} \cdot 1 = 0 \\ \dfrac{\partial I}{\partial p_1} = \sum_{n=1}^{N} 2 \cdot \{p_0 + p_1 \cdot x_n - y_n\} \cdot x_n = 0 \end{cases}$$

3. K=2 の例

$$\begin{bmatrix} p_0 \\ p_1 \end{bmatrix} = \begin{bmatrix} N & \sum_{n=1}^{N} x_n \\ \sum_{n=1}^{N} x_n & \sum_{n=1}^{N} x_n^2 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \sum_{n=1}^{N} y_n \\ \sum_{n=1}^{N} y_n x_n \end{bmatrix}$$

最小自乗法（詳細）

$$I = \sum_{i=0}^{N-1} \{y(x_i) - y_i\}^2 = \sum_{i=0}^{N-1} \{a + b x_i - y_i\}^2 \qquad \leftarrow \quad y(x_i) = a + b x_i$$

より,

$$\frac{\partial I}{\partial a} = \frac{\partial}{\partial a} \sum_{i=0}^{N-1} \{a + b x_i - y_i\}^2 = 2 \sum_{i=0}^{N-1} \{a + b x_i - y_i\} = 0 \qquad \leftarrow \quad \frac{df(g(x))}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx}$$

$$\frac{\partial I}{\partial b} = \frac{\partial}{\partial b} \sum_{i=0}^{N-1} \{a + b x_i - y_i\}^2 = 2 \sum_{i=0}^{N-1} \{a + b x_i - y_i\} x_i = 0$$

すなわち,

$$\begin{pmatrix} \sum_{i=0}^{N-1} 1 & \sum_{i=0}^{N-1} x_i \\ \sum_{i=0}^{N-1} x_i & \sum_{i=0}^{N-1} x_i^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^{N-1} y_i \\ \sum_{i=0}^{N-1} x_i y_i \end{pmatrix}$$

以上より,

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{\sum x_i^2 \sum 1 - \sum x_i \sum x_i} \begin{pmatrix} \sum x_i^2 & -\sum x_i \\ -\sum x_i & \sum 1 \end{pmatrix} \begin{pmatrix} \sum y_i \\ \sum y_i x_i \end{pmatrix}$$

方法2 【一般化逆行列】

行列で表す

$$\underset{\mathbf{Y}}{\begin{bmatrix}1\\0\\2\end{bmatrix}} = \underset{\mathbf{X}}{\begin{bmatrix}1 & 0\\1 & 1\\1 & 2\end{bmatrix}} \underset{\mathbf{P}}{\begin{bmatrix}p_0\\p_1\end{bmatrix}}$$

データ　直線上の点

n	x_n	y_n		$f(x)$
1	0	1	=	$p_0+p_1 \cdot 0$
2	1	0	=	$p_0+p_1 \cdot 1$
3	2	2	=	$p_0+p_1 \cdot 2$

↑ 等しいとおく...

方程式を解く

$\mathbf{P} = \mathbf{X}^{-1}\,\mathbf{Y}$ ... ではなくて

$\mathbf{P} = \mathbf{X}^{+}\,\mathbf{Y}$ ... で求まる

ここで、

$\mathbf{X}^{+} = (\mathbf{X}^{T}\,\mathbf{X})^{-1}\,\mathbf{X}^{T}$ 一般化逆行列

$$\mathbf{X}^{+} = \left(\begin{bmatrix}1 & 1 & 1\\0 & 1 & 2\end{bmatrix} \cdot \begin{bmatrix}1 & 0\\1 & 1\\1 & 2\end{bmatrix} \right)^{-1} \cdot \begin{bmatrix}1 & 1 & 1\\0 & 1 & 2\end{bmatrix}$$

$$\begin{bmatrix}p_0\\p_1\end{bmatrix} = \mathbf{X}^{+} \begin{bmatrix}1\\0\\2\end{bmatrix} = \begin{bmatrix}0.5\\0.5\end{bmatrix}$$

一般化逆行列は　2種類

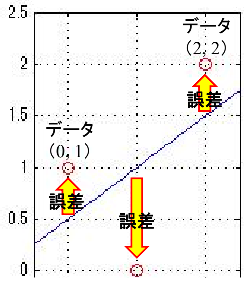
$$\mathbf{X}^{+} = (\mathbf{X}^{T} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^{T}$$

縦長な X に使う（方程式が変数より 多い・優決定）
最小自乗の意味で最適

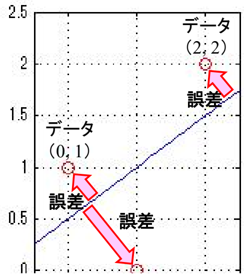
$$\mathbf{X}^{+} = \mathbf{X}^{T} \cdot (\mathbf{X} \cdot \mathbf{X}^{T})^{-1}$$

横長な X に使う（方程式が変数より少ない・劣決定）
$\|X\|_2$ が最小

誤差の測り方は　2通り



$$\min \sum_{n=1}^{3} \{ f(x_n) - y_n \}^2$$



$$\min \sum_{n=1}^{3} \{ (x - x_n)^2 + (y - y_n)^2 \}$$
$$s.t. \quad y = f(x)$$

ラグランジュの未定乗数法

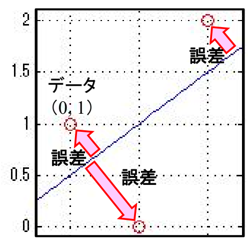
$f(x) = p_0 + p_1 \cdot x$ については,

$$\frac{\partial J}{\partial p_0} = 0, \quad \frac{\partial J}{\partial p_1} = 0$$

に対して,

$$\frac{\partial J}{\partial \lambda} = 0$$

を加えて解けばよい,



$$J = \sum_{n=1}^{3} \{ (x - x_n)^2 + (y - y_n)^2 \} + \lambda \ \{ y - f(x) \}$$

$$\min \sum_{n=1}^{3} \{ (x - x_n)^2 + (y - y_n)^2 \}$$
$$s.t. \quad y - f(x) = 0$$

拘束条件付きの　最小二乗法

$f(x) = \sum_{k=1}^{K} p_k \cdot x^k$ については,

$$\frac{\partial J}{\partial p_i} = 0, \quad i = 1, 2, \cdots$$

に対して,

$$\frac{\partial J}{\partial \lambda} = 0$$

を加えて解けばよい,

λ: ラグランジュの未定乗数

$$J = \sum_{n=1}^{3} \{ (x - x_n)^2 + (y - y_n)^2 \} + \lambda \ \{ y - f(x) \}$$ ← 拘束条件

$$\min \sum_{n=1}^{3} \{ (x - x_n)^2 + (y - y_n)^2 \}$$
$$s.t. \quad y - f(x) = 0$$ ← 拘束条件