



HOKKAIDO UNIVERSITY

Title	2005年度 情報理論講義ノート
Author(s)	井上, 純一; Inoue, Jun-ichi
Description	http://www005.upp.so-net.ne.jp/j_inoue/index.html http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/
Issue Date	2005-11-18T09:19:52Z
Doc URL	https://hdl.handle.net/2115/772
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/2.1/jp/
Type	learning object
File Information	InfoTheory05_9.pdf, 第9回講義ノート



情報理論 配布資料 #9

担当：井上 純一 (情報科学研究科棟 8-13)

URL : http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/

平成 17 年 6 月 20 日

目次

7.5	線形符号	56
7.5.1	2 進線形符号	56
7.5.2	パリティ検査行列	57
7.5.3	パリティ検査行列と最小距離	58
7.5.4	線形符号の符号化	59
7.5.5	線形符号の復号化	59

演習問題 8 の解答例

この通信路では 2 元対称通信路と異なり, 入力空間と出力空間の対称性が無いところに注意する.

まずは 2 元対称消失通信路の通信路容量を計算しておく. 入力のある特定の値に固定した場合の条件付きエントロピーは

$$\begin{aligned} H(Y|X=0) &= - \sum_{y=0,1,x} P_{Y|X}(y|0) \log P_{X|Y}(y|0) \\ &= -p \log p - q \log q - (1-p-q) \log(1-p-q) = H(Y|X=1) \equiv h(p, q) \end{aligned} \quad (233)$$

となるので, 条件付きエントロピーは

$$H(Y|X) = \sum_{x=0,1} P_X(x) H(Y|X=x) = h(p, q) \sum_{x=0,1} P_X(x) = h(p, q) \quad (234)$$

であり, 2 元対称通信路の場合と同様に入力分布 $P_X(x)$ に依らない. 従って, 通信路容量を求めるには, 出力のエントロピーを入力分布に対して最大化すればよい. 例によって, $P_X(0) = t, P_X(1) = 1-t$ とおくと

$$P_Y(0) = (1-p-q)t + p(1-t) \quad (235)$$

$$P_Y(1) = pt + (1-p-q)(1-t) \quad (236)$$

$$P_Y(x) = q \quad (237)$$

であるから

$$\begin{aligned} H(Y) &= -[(1-p-q)t + p(1-t)] \log[(1-p-q)t + p(1-t)] \\ &\quad - [pt + (1-p-q)(1-t)] \log[pt + (1-p-q)(1-t)] \\ &\quad - q \log q \end{aligned} \quad (238)$$

を最大化する t の値は $t = 1/2$ であり、このときの最大値は

$$\max_t H(Y) = 1 - q - (1 - q) \log(1 - q) - q \log q \quad (239)$$

であることがわかる。従って、通信路容量は

$$C = 1 - q - (1 - q) \log(1 - q) - q \log q - h(p, q) \quad (240)$$

で与えられる。ここで、この通信路は $q = 0$ で 2 元対称通信路と一致するが、上式で $q = 0$ と置けば、前回学んだ 2 元対称通信路の通信路容量： $C = 1 - h(p)$ となることは確認しておくべきであろう。

さて、次に転送によって受信系列の広がる大きさ、通信路出力の典型列の大きさ w を評価しておこう。これは

$$w = {}_n C_{np} \times {}_{n(1-p)} C_{nq} \quad (241)$$

で与えられるので、この両辺の対数をとって、スターリングの公式： $\log n! \simeq n \log n - n$ を用いると直ちに

$$\log w = n \{-p \log p - q \log q - (1 - p - q) \log(1 - p - q)\} = nh(p, q) \quad (242)$$

すなわち

$$w = 2^{nh(p, q)} \quad (243)$$

のように評価できる。これは先ほどの $h(p, q)$ の導出と重ね合わせて考えると、 $w = 2^{H(Y|X)}$ のように書き直すことができる。

ところで今までとは逆に、ある一つの受信系列を受取ったときに、それはどの程度の範囲の送信系列からやってきたものとみなせるであろうか？ (図 26 参照) この答えは $w' = 2^{H(X|Y)}$ で与えられる。前回見た 2

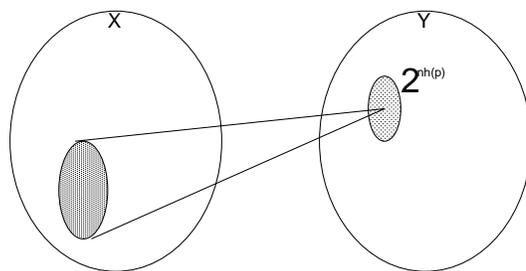


図 26: ある一つの受信系列を受取ったときに、それはどの程度の範囲の送信系列からやってきたものか、を評価する。

元対称通信路の場合には、この大きさ w' は $H(X|Y) = H(Y|X) = h(p)$ で与えられたので、 $w' = w$ であった。しかし、今の場合入力と出力の間の対称性は無いので、 $H(X|Y)$ をまじめに評価しなければならない。しかし、これはさほど難しく無く

$$\begin{aligned} H(X|Y) &= H(X) - H(Y) + H(Y|X) \\ &= n \{1 - (1 - q) + (1 - q) \log(1 - q) + q \log q + h(p, q)\} \end{aligned} \quad (244)$$

のように求めることができる。従って、復号によって誤りが生じるのは、入力の空間において $(M - 1)/2^n$ の確率で選び出された該当符号語以外の符号語が選び出され、これが $w' = 2^{H(X|Y)}$ の大きさの領域に落ち

る場合であるから

$$\begin{aligned} P_E &= \frac{M-1}{2^n} \times 2^{n\{1-(1-q)+(1-q)\log(1-q)+q\log q+h(p,q)\}} \\ &\simeq 2^{n\{R-((1-q)-(1-q)\log(1-q)-q\log q-h(p,q))\}} = 2^{n(R-C)} \end{aligned} \quad (245)$$

$$C = 1 - q - (1 - q) \log(1 - q) - q \log q - h(p, q) \quad (246)$$

つまり、 $R < C$ であれば、 $n \rightarrow \infty$ で $P_E \rightarrow 0$ となる（もちろん、伝送速度の定義はこの 2 元対称消失通信路でも変わらないので、 M も変化せずに $M = 2^{nR}$ である）。これで (i) の証明は終わった。

次は (ii) であるが、これは簡単で、この通信路の場合の箱のサイズ z は

$$z = \frac{n C_{nq} \times n(1-q) C_{n(1-q)/2}}{M} \quad (247)$$

であるから、例に $n \rightarrow \infty$ の場合にスターリングの公式を用いて分子を評価すれば、簡単に

$$z = \frac{2^{n\{(1-q)-(1-q)\log(1-q)-q\log q\}}}{M} \quad (248)$$

となる。ここで、上式で $q = 0$ とすると、前回見た 2 元対称通信路の場合の「箱」のサイズ $z = 2^n/M$ となることを確認しておくべきである。

さて、 $z > w$ でなければならないので、この条件を書き下してみると

$$2^{n(R-C)} < 1 \quad (249)$$

であり、明らかに $R > C$ では成立しない。従って、(ii) が証明された。

7.5 線形符号

ここからは具体的な符号化と復号化法について見ていく。

7.5.1 2進線形符号

次の 2 つの条件を満たす $\{0, 1\}^n$ の線形部分空間 C を 2 進線形符号と呼ぶ。

2 進線形符号の条件：

- (1) $x, y \in C$ ならば $x + y \in C$
- (2) $\lambda \in \{0, 1\}$, $x \in C$ ならば $\lambda x \in C$

C は $\{0, 1\}^n$ の部分空間なので、独立な基底ベクトル $x_1, x_2, \dots, x_k \in C$ ($k \leq n$) を選べる。よって、 C は

$$\begin{aligned} C &= \{u_1 x_1 + \dots + u_k x_k \mid u_i \in \{0, 1\}, i = 1, \dots, k\} \\ &= \{uG \mid u = \{u_1, \dots, u_k\} \in \{0, 1\}^k\} \end{aligned} \quad (250)$$

のように書ける。ここで、 G は x_i ($i = 1, \dots, k$) を縦にならべた行列：

$$G = \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_k \end{pmatrix} \quad (251)$$

であり、生成行列と呼ばれる。

(例)

$$C = \left\{ \begin{array}{l} (00000), (01100), (00110), (01010) \\ (11111), (10011), (11001), (10101) \end{array} \right\} \quad (252)$$

とすれば、 C のどの元どうしの和も C に属しているの、線形符号の条件 (1) を満たしている。条件 (2) を満たすことも明らか。

どの 2 つの和でももう一つを表せない基底ベクトルとして $(01100), (00110), (11111)$ をとれる。この基底ベクトル間の最小距離は 2 である。従って、生成行列は

$$G = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \quad (253)$$

ここで、この行列の行を $1 \rightarrow 3, 2 \rightarrow 1, 2 \rightarrow 3$ のように並びかえ、その行列の 1 行目と 2 行目、2 行目と 3 行目を足す基本変形を行うと

$$G^* = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} = (I_3 : G_1) \quad (254)$$

と書き直すことができる。ここで

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad G_1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} \quad (255)$$

である。

7.5.2 パリティ検査行列

今後、 A' のように記号の右肩に「ダッシュ」「プライム」をつけることによって行列の転値を表すことにする。このとき、線形符号 C はパリティ検査行列と呼ばれる行列 H を用いて

$$C = \{ \mathbf{x} \in \{0,1\}^n \mid H\mathbf{x}' = \mathbf{0}' \} \quad (256)$$

のように表すことができる。 H はそのランクが $\text{rank}(H) = n - k$ 、サイズが $(n - k) \times n$ の行列である。

\mathbf{x}_i ($i = 1, \dots, k$) に対しても $H\mathbf{x}'_i = 0$ が成り立つので、明らかに

$$GH' = 0 \quad (257)$$

が成り立つ。特に、 $G^* = (I_k, G_1)$ のように書き換えた生成行列に対し、パリティ検査行列も $H^* = (G_1' : I_{n-k})$ と書き換えることができる

$$G^* H^{*'} = (I_k : G_1) \begin{pmatrix} G_1 \\ I_{n-k} \end{pmatrix} = I_k G_1 + G_1 I_{n-k} = G_1 + G_1 = 0 \quad (258)$$

が成り立つ。

(例)

$$G^* = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} \quad (259)$$

に対して

$$H^* = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (260)$$

であり、このとき確かに

$$G^* H^{*'} = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = 0 \quad (261)$$

が成立する。

このパリティ検査行列に対し、 $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)$ とすると

$$H^* \mathbf{x}' = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = 0 \quad (262)$$

より、 x_4, x_5 は x_1, x_2, x_3 を用いて

$$x_4 = x_1 + x_2 + x_3 \quad (263)$$

$$x_5 = x_1 \quad (264)$$

と書ける。従って、 $(x_1, x_2, x_3) = (000)(001)(010)(100)(011)(101)(110)(111)$ と選ぶと $(x_1, x_2, x_3, x_4, x_5) = (00000)(00110)(10110)(10011)(01100)(10101)(11001)(11111)$ のように線形符号 C が得られる。

7.5.3 パリティ検査行列と最小距離

パリティ検査行列：

$$H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix} \quad (265)$$

に対して、 $H\mathbf{x}' = 0$ は

$$x_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + x_5 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 0 \quad (266)$$

と書き下すことができる。この方程式の自明な解は $x = (00000)$ であり、この解のハミング重み、つまり、ゼロでないビット数は $w(x) = 0$ である。 $w(x) = 1, 2$ となる (266) 式の解はなく、 $w(x) = 3$ となる x は $x = (11100)$ である (C の最小重み (最小距離) は 3 である)。

定理 5・3 :

線形符号のパリティ検査行列 $H = (h_1, \dots, h_n)$ のどの t 個の列ベクトルも線形独立で、ある $t+1$ 個の列ベクトルが線形従属のとき、 C の最小距離 d は $t+1$ である。

7.5.4 線形符号の符号化

パリティ検査行列 :

$$H = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}, \quad C = \{x \in \{0, 1\}^5 \mid Hx' = \mathbf{0}'\} \quad (267)$$

に対し、 $Hx' = \mathbf{0}'$ を書き出してみると

$$x_3 = x_1 \quad (268)$$

$$x_4 = x_1 + x_2 \quad (269)$$

$$x_5 = x_2 \quad (270)$$

となり、 x_1, x_2 (情報ビット) を決めると、 x_3, x_4, x_5 (パリティ検査ビット) が決まる。 $(x_1, x_2) = (00)(01)(10)(11)$ とすると $(x_3, x_4, x_5) = (000)(011)(110)(101)$ なので、線形符号は

$$C = \{(00000), (01011), (10110), (11101)\} \quad (271)$$

のように作ることができる。

7.5.5 線形符号の復号化

受信データを y 、送信データを x 、誤りベクトルを e とすると

$$y = x + e \quad (272)$$

であり、この両辺に左からパリティ検査行列 H をかけると

$$Hy' = Hx' + He' = He' \quad (273)$$

となる。ここで、 $s' = He'$ を誤りベクトル (コセットリーダー) e のシンδροームと呼ぶ。

このとき、線形符号の復号化法としては、 $s' = Hy'$ を計算し、この s' に対応するコセットリーダーを y に加えれば良い。

次の例題を見てみよう。

例題 9

成分が $0, 1$ である任意のベクトル $\mathbf{u} = (u_1, u_2, \dots, u_n), \mathbf{v} = (v_1, v_2, \dots, v_n)$ に対し, 和を $\mathbf{u} + \mathbf{v} = (u_1 \oplus v_1, u_2 \oplus v_2, \dots, u_n \oplus v_n)$ で定義する. ここで, 記号 \oplus は排他的論理和を表し, 交換則 $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ が成り立つ. また, ベクトルにスカラー $a \in \{0, 1\}$ をかける演算を $a\mathbf{u} = (au_1, au_2, \dots, au_n)$ で定義し, 2つのスカラーの積は $00 = 01 = 10 = 0, 11 = 1$ とする. つまり, $a = 0$ のとき $a\mathbf{u} = \mathbf{0}$, $a = 1$ のとき $a\mathbf{u} = \mathbf{u}$ である. ここに $\mathbf{0} = (0, 0, 0, \dots, 0)$ はゼロベクトルである. これらの演算が成り立つベクトルに対して以下の問に答えよ.

- (1) 任意のベクトル \mathbf{v} に対し, $\mathbf{v} + \mathbf{e} = \mathbf{v}$ となるような \mathbf{e} は何か.
- (2) 任意のベクトル \mathbf{v} に対し, $\mathbf{v} + \mathbf{f} = \mathbf{0}$ となるような \mathbf{f} は何か.
- (3) 任意のベクトル \mathbf{u} の成分中 $u_i \neq 0$ となる i の個数をハミング重みと呼び $w(\mathbf{u})$ で表す. このとき任意のベクトル \mathbf{u}, \mathbf{v} に対し, 不等式:

$$w(\mathbf{u}) + w(\mathbf{v}) \geq w(\mathbf{u} + \mathbf{v})$$

が成り立つことを示せ.

- (4) パリティ検査行列:

$$\mathbf{H} = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

を持つ線形符号に対し, 1ビット誤りベクトルの復号表を作成せよ. つまり, 1ビット誤りベクトル $\mathbf{z} = (1, 0, 0, 0, 0, 0), (0, 1, 0, 0, 0, 0), (0, 0, 1, 0, 0, 0), \dots, (0, 0, 0, 0, 0, 1)$ に対して, \mathbf{z} と $\mathbf{s}_z \equiv \mathbf{z}\mathbf{H}^T$ の対応表を作れ. ここに \mathbf{H}^T は \mathbf{H} の転置である. また, この復号表から, この線形符号が 1ビット誤り訂正符号であるか判定せよ.

- (5) 送信ベクトル \mathbf{u} が通信路を通過後, 1ビットノイズ \mathbf{z} が加わり, $\mathbf{r} = \mathbf{u} + \mathbf{z}$ として受信された. 受信ベクトル \mathbf{r} に対し, $\mathbf{s}_r \equiv \mathbf{r}\mathbf{H}^T$ を計算し, (4) で作成した復号表から $\mathbf{s}_z = \mathbf{s}_r$ となる \mathbf{z} を求めると, 送信ベクトルは $\mathbf{u} = \mathbf{r} + \mathbf{z}$ として復元できる. この方法を用いて受信ベクトルが $\mathbf{r} = (0, 1, 0, 1, 1, 0)$ のとき, \mathbf{u} 及び \mathbf{z} を求めよ.
- (6) この符号を誤り率 p の 2元対称通信路で用いたときの誤り確率 p_z を求めよ.

(解答例)

- (1) $\mathbf{e} = \mathbf{0}$.
- (2) $\mathbf{f} = \mathbf{v}$.
- (3) 任意のベクトル $\mathbf{x}, \mathbf{y}, \mathbf{z}$ に対し, ハミング距離 d は次の三角不等式:

$$d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}) \geq d(\mathbf{x}, \mathbf{z}) \quad (274)$$

を満たす. ここで

$$d(\mathbf{x}, \mathbf{y}) = w(\mathbf{x} - \mathbf{y})$$

$$= w(\mathbf{x} + \mathbf{y}) \quad (275)$$

$$\begin{aligned} d(\mathbf{y}, \mathbf{z}) &= w(\mathbf{y} - \mathbf{z}) \\ &= w(\mathbf{y} + \mathbf{z}) \end{aligned} \quad (276)$$

$$\begin{aligned} d(\mathbf{x}, \mathbf{z}) &= w(\mathbf{x} - \mathbf{z}) \\ &= w(\mathbf{x} + \mathbf{z}) \end{aligned} \quad (277)$$

$$(278)$$

が成り立つ。従って (1) 式の三角不等式はハミング重みを用いて

$$w(\mathbf{x} + \mathbf{y}) + w(\mathbf{y} + \mathbf{z}) \geq w(\mathbf{x} + \mathbf{z}) \quad (279)$$

と書き直すことができる。ところで

$$\mathbf{x} + \mathbf{y} = \mathbf{u} \quad (280)$$

$$\mathbf{y} + \mathbf{z} = \mathbf{v} \quad (281)$$

とおき, (7)(8) 式の辺々を足してみると

$$\mathbf{x} + \mathbf{z} + (\mathbf{y} + \mathbf{y}) = \mathbf{u} + \mathbf{v} \quad (282)$$

つまり

$$\mathbf{x} + \mathbf{z} = \mathbf{u} + \mathbf{v} \quad (283)$$

が得られるから, 結局 (6) の不等式は

$$w(\mathbf{u}) + w(\mathbf{v}) \geq w(\mathbf{u} + \mathbf{v}) \quad (284)$$

となり, 題意が示された.

(4) 例えば 1 ビット誤りベクトル $\mathbf{z} = (1, 0, 0, 0, 0, 0)$ に対し, $\mathbf{s}_z = \mathbf{z}\mathbf{H}^T$ を計算すると

$$\begin{aligned} \mathbf{s}_z &= \mathbf{z}\mathbf{H}^T \\ &= (1, 0, 0, 0, 0, 0) \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= (0, 1, 1) \end{aligned} \quad (285)$$

となるから, 1 ビット誤りベクトルの復号表は

\mathbf{z} (1 ビット誤りベクトル)	$\mathbf{s}_z = \mathbf{z}\mathbf{H}^T$
(1,0,0,0,0,0)	(0,1,1)
(0,1,0,0,0,0)	(1,0,1)
(0,0,1,0,0,0)	(1,1,0)
(0,0,0,1,0,0)	(1,0,0)
(0,0,0,0,1,0)	(0,1,0)
(0,0,0,0,0,1)	(0,0,1)

となる。また、上記表から全ての 1 ビット誤りベクトルには異なる s_z が対応しているので、この線形符号は 1 ビット誤り訂正符号である。

(5) 送信ベクトル $r = (0, 1, 0, 1, 1, 0)$ に対し $s_r = rH^T$ は

$$\begin{aligned} s_r &= rH^T \\ &= (0, 1, 0, 1, 1, 0) \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= (0, 1, 1) \end{aligned} \quad (286)$$

となる。従って、問題 (4) で求めた復号表から該当する 1 ビット誤りベクトルを探すと、これは $z = (1, 0, 0, 0, 0, 0)$ であり、送信ベクトルの 1 ビット目に誤りがあることがわかる。また、送信ベクトル u は

$$\begin{aligned} u &= r + z \\ &= (1, 1, 0, 1, 1, 0) \end{aligned} \quad (287)$$

である。よって求める答えは

$$u = (1, 1, 0, 1, 1, 0) \quad (288)$$

$$z = (1, 0, 0, 0, 0, 0) \quad (289)$$

である。

(6) 誤りなし確率は全ビットが誤りなしに送られる場合と、1 ビット誤りが全 6 ビットの中の何処か 1 箇所にある場合であるから、これと余事象の確率を考えれば誤り確率 p_z は

$$p_z = 1 - (1 - p)^6 - 6p(1 - p)^5 \quad (290)$$

となる。

演習問題 9

定理 5・3 を証明せよ。