

ラグランジュ乗数法による制約条件付きの解

前川眞一

20071113,20140818,20161031,20210701,20220602 2022. 6. 2

1 線型モデルの母数への線形制約

y, X, β をそれぞれ $n \times 1, n \times q, q \times 1$ の行列として、

$$y = X\beta + \varepsilon = \hat{y} + \varepsilon \quad (1)$$

$$\hat{y} = X\beta \quad (2)$$

というモデルがあった場合、最小2乗基準

$$\text{RSS} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = (y - \hat{y})'(y - \hat{y}) \quad (3)$$

を最小とする β , すなわち $\hat{\beta}$ は、

$$\hat{\beta} = (X'X)^{-1}X'y \quad (4)$$

で与えられるが、RSS を r 個の線形制約条件

$$L\beta = c \quad (5)$$

の下で最小化したい場合を考える。ただし、 L ならびに c は、それぞれ、 $r \times q, r \times 1$ の行列であり、 $\text{rank}(L) = \text{nrows}(L) = r < q$ とする。

制約条件の例としては、 β を 4×1 のベクトルとして、 $\beta_1 = \beta_2$, & $\beta_3 = \beta_4$ 等の線形期待値構造の制約条件があるが、これは以下に示すような階数 2 の行列

$$L = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \quad \text{and} \quad c = \mathbf{0} \quad (6)$$

を用いて書き表すことが出来る。

1.1 ラグランジュ乗数法による解

m を $r \times 1$ の未知のベクトルとして、

$$\text{RSS1} = (y - X\beta)'(y - X\beta) + 2m'(L\beta - c) \quad (7)$$

$$= y'y - 2y'X\beta + \beta X'X\beta + 2m' L\beta - 2m'c \quad (8)$$

を β と m の両方に関して最小化して解を求める方法を線形制約のためのラグランジュの未定乗数法と呼ぶ。なお、この m がラグランジュの未定乗数 (ベクトル) と呼ばれるものである。

Eq.(7) の β と m に関する偏微分を 0 と置いた方程式は

$$\frac{\partial \text{RSS1}}{\partial \beta} = 2(X'X\beta - X'y + L'm) = \mathbf{0} \quad (9)$$

$$\frac{\partial \text{RSS1}}{\partial m} = L\beta - c = \mathbf{0} \quad (10)$$

となり、制約付きの解はこの連立方程式の解となっている。なお、未定乗数に関する偏微分部分、Eq.(10) に制約式そのものが出る。

Eq.(9) を、 β に関して解くと、

$$\beta = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{y} - \mathbf{L}'\mathbf{m}) \quad (11)$$

を得るが、ここで、式の簡略化のため

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (12)$$

を用いると、

$$\beta = \hat{\beta} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}'\mathbf{m} \quad (13)$$

となる。この両辺の左から \mathbf{L} を掛けて制約条件を用いると、

$$\mathbf{L}\beta = \mathbf{L}\hat{\beta} - \mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}'\mathbf{m} = \mathbf{c} \quad (14)$$

を得るが、 \mathbf{L} は行フルランクなので、 \mathbf{m} は

$$\mathbf{m} = (\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}')^{-1}(\mathbf{L}\hat{\beta} - \mathbf{c}) \quad (15)$$

となる。そして、これを Eq.(13) に代入すると以下の制約付きの回帰係数の解を得る。

$$\beta = \hat{\beta} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}'\mathbf{m} \quad (16)$$

$$= \hat{\beta} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}'(\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}')^{-1}(\mathbf{L}\hat{\beta} - \mathbf{c}) \quad (17)$$

そして、この解が確かに制約条件を満たすことは、左から \mathbf{L} を掛けて確かめることができる。

$$\mathbf{L}\beta = \mathbf{L}\hat{\beta} - \mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}'(\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}')^{-1}(\mathbf{L}\hat{\beta} - \mathbf{c}) \quad (18)$$

$$= \mathbf{L}\hat{\beta} - (\mathbf{L}\hat{\beta} - \mathbf{c}) = \mathbf{c} \quad (19)$$

なお、このような線形期待値構造の制約条件を満たす解は、後述する双対系を用いて求めることも出来る。

2 ラグランジュの未定乗数法とその例

一般的に、 p 次元の変数ベクトル β のスカラー関数 $f(\beta)$ を r 個の制約条件

$$g_1(\beta) = c_1, \quad g_2(\beta) = c_2, \quad \dots, \quad g_r(\beta) = c_r \quad (20)$$

の元で最小化もしくは最大化する β は、新たな目的関数を

$$q(\beta, \mathbf{m}) = f(\beta) + \mathbf{m}'(\mathbf{g}(\beta) - \mathbf{c}) \quad (21)$$

とし、その β と \mathbf{m} に関する偏微分を $\mathbf{0}$ と置いて得られる連立方程式

$$\frac{\partial q(\beta, \mathbf{m})}{\partial \beta} = \frac{\partial f(\beta) + \mathbf{m}'(\mathbf{g}(\beta) - \mathbf{c})}{\partial \beta} = \mathbf{0} \quad (22)$$

$$\frac{\partial q(\beta, \mathbf{m})}{\partial \mathbf{m}} = \frac{\partial f(\beta) + \mathbf{m}'(\mathbf{g}(\beta) - \mathbf{c})}{\partial \mathbf{m}} = \mathbf{0} \quad (23)$$

の解として与えられる。ただし $g(\beta)$ は上記の制約条件を縦に並べた $r \times 1$ のベクトル、また m は $r \times 1$ のラグランジュの未定乗数ベクトルである。なお、上記の方程式は $r \times p$ の $g(\beta)$ の Jacobian 行列 $\left(\frac{\partial g(\beta)}{\partial \beta}\right)$ を使って

$$\frac{\partial f(\beta)}{\partial \beta} + \left(\frac{\partial g(\beta)}{\partial \beta}\right)' m = 0 \quad \text{and} \quad g(\beta) - c = 0 \quad (24)$$

と書けるが、 m に関する偏微分の結果から条件式そのものが出ている。

以下にいくつかの例を示すが、変数としては β では無く x や y 等の様々なものが使われている。またラグランジュの乗数としても λ や L が用いられている。

2.1 例

$$f(x, y) = x^2 + y^2 - 3x - 4y \quad (25)$$

は、 $x = 1.5, y = 2$ において最小値 -6.25 を取る。この関数を、制約条件

$$x - y = 1 \quad (26)$$

の下で最小化する。

元の関数と、制約条件に未定常数 λ を掛けたものを x, y, λ の関数と見なし、

$$g(x, y, \lambda) = f(x, y) + \lambda(x - y - 1) \quad (27)$$

その制約無しでの最小値を求める。

g の各変数に関する偏微分を 0 とおいた方程式は

$$\frac{\partial g}{\partial x} = 2x + \lambda - 3 = 0 \quad (28)$$

$$\frac{\partial g}{\partial y} = 2y - \lambda - 4 = 0 \quad (29)$$

$$\frac{\partial g}{\partial \lambda} = x - y - 1 = 0 \quad (30)$$

となる。(最小とすべき関数 g のラグランジュの乗数に関する偏微分は、必ず制約条件式となる。)

上のふたつから、

$$x = (3 - \lambda)/2 \quad \text{and} \quad y = (4 + \lambda)/2 \quad (31)$$

を得るが、これを 3 番目、すなわち制約条件そのものに代入すると

$$-1 - \frac{1}{2}(\lambda + 4) - \frac{\lambda - 3}{2} = 0 \quad (32)$$

を得る。これから

$$\lambda = -\frac{3}{2} \quad (33)$$

となり、これを上式に代入すると

$$x = \frac{9}{4} \quad \text{and} \quad y = \frac{5}{4} \quad (34)$$

を得る。また、そのときの最小値は $f = -\frac{41}{8}$ となる。

2.2 例

例 1 と同じ関数

$$f(x, y) = x^2 + y^2 - 3x - 4y \quad (35)$$

を、異なる制約条件

$$x^2 + y^2 = 25 \quad (36)$$

の下で最小化する。

元の関数と、制約条件に未定常数 λ を掛けたものを x, y, λ の関数と見なし、

$$g(x, y, \lambda) = f(x, y) + \lambda(x^2 + y^2 - 25) \quad (37)$$

その制約無しでの最小値を求める。

g の各変数に関する偏微分を 0 とおいた方程式は

$$\frac{\partial g}{\partial x} = 2x + 2\lambda x - 3 = 0 \quad (38)$$

$$\frac{\partial g}{\partial y} = 2y + 2\lambda y - 4 = 0 \quad (39)$$

$$\frac{\partial g}{\partial \lambda} = x^2 + y^2 - 25 = 0 \quad (40)$$

となる。(最小とすべき関数 g のラグランジュの乗数に関する偏微分は、必ず制約条件式となる。)

上のふたつから、

$$x = \frac{3}{2(\lambda + 1)} \quad \text{and} \quad y = \frac{2}{\lambda + 1} \quad (41)$$

を得るが、これを 3 番目、すなわち制約条件そのものに代入すると

$$\frac{25}{4(\lambda + 1)^2} - 25 = 0 \quad (42)$$

を得る。これから

$$\lambda = -\frac{3}{2} \quad \text{or} \quad \lambda = -\frac{1}{2} \quad (43)$$

となり、これを上式に代入すると

$$x = -3, \quad y = -4 \quad \text{or} \quad x = 3, \quad y = 4 \quad (44)$$

を得る。また、そのときの f の最小値は、 $(x=3, y=4)$ において 0 となる。(最大値は 50 となる。)

2.3 例

直交プロクラスティス回転

$$Q = \text{tr} \left[(C - AT)'(C - AT) \right] \quad (45)$$

を最小とする直交行列 T を求める。ただし、 $T'T = I$ である。

$$Q' = \text{tr} \left[(C - AT)'(C - AT) \right] + \text{tr} (L(T'T - I)) \quad (46)$$

とする。ただし L はラグランジュの未定乗数からなる行列である。

$$\frac{\partial Q'}{\partial T} = -2A'C + 2AA'T + T(L + L') \quad (47)$$

より

$$\frac{1}{2}T(L + L') = A'C - A'AT \quad (48)$$

を得るが、この両辺の左から T' を掛け条件を使うと

$$\frac{1}{2}(L + L') = T'(A'C - A'AT) = T'A'C - T'A'AT \quad (49)$$

を得る。しかし、 $(L + L')$ 行列は対称であるため、

$$T'A'C - T'A'AT = (T'A'C - T'A'AT)' \quad (50)$$

でなければならず、これより

$$T'A'C = C'AT \quad (51)$$

を得るが、上式を満たす T 行列は以下のようにして得られる。すなわち、 $A'C$ の特異値分解を

$$A'C = UDV \quad (52)$$

とすれば、

$$T = UV \quad (53)$$

である。なお、この T は $A'CC'A$ のスペクトル分解（固有値分解）を

$$A'CC'A = UD^2U' \quad (54)$$

としたときに

$$T = UD^{-1}U'A'C \quad (55)$$

と書くことも出来る。

2.4 例

線形回帰の残差の2乗和

$$\text{RSS} = \sum_{i=1}^n \left(v_i - \sum_{j=1}^q u_{ij}\xi_j \right)^2 = (v - U\xi)'(v - U\xi) \quad (56)$$

を

$$\|U\xi\|^2 = d$$

という制約の下で RSS を最小とする解 ξ^{**} を求める。ただし、RSS の制約無しの最小値を与える解を

$$\hat{\xi} = (U'U)^{-1}U'v \quad (57)$$

とする。ラグランジュの乗数を λ とすれば、制約条件下での最小化基準は

$$Q = \text{RSS} + \lambda(d - \xi'U'U\xi)$$

となり、その偏微分は以下ようになる。

$$\begin{aligned} \frac{\partial Q}{\partial \xi} &= -2U'v + 2U'U\xi - 2U'U\xi \\ &= 2(-U'v + (1 - \lambda)U'U\xi) \end{aligned}$$

したがって、

$$\begin{aligned} \xi^{**} &= \frac{1}{(1 - \lambda)}(U'U)^{-1}U'v \\ &= \frac{1}{(1 - \lambda)}\hat{\xi} \end{aligned}$$

ここで、 Q のラグランジュの乗数に関する偏微分を 0 と置いて得られる制約条件式に上記の ξ^{**} を代入して得られる

$$\begin{aligned} \|U\xi^{**}\|^2 &= \left(\frac{1}{(1 - \lambda)}\right)^2 \hat{\xi}'U'U\hat{\xi} \\ &= \left(\frac{1}{(1 - \lambda)}\right)^2 \hat{v}'\hat{v} \\ &= d \end{aligned}$$

の関係を使うと

$$\frac{1}{(1 - \lambda)} = \sqrt{\frac{d}{\hat{v}'\hat{v}}}$$

すなわち、

$$\xi^{**} = \frac{\sqrt{d}}{\|\hat{v}\|}\hat{\xi}$$

となる。

2.5 例

p カテゴリの多項分布の確率を $\pi_j, j = 1, 2, \dots, p,$ で表すと、

$$\sum_{j=1}^p \pi_j = 1 \quad (58)$$

である。第 j 番目のカテゴリの度数を f_j とした時のに、Eq.(58) の条件の下で対数尤度

$$\ln L = \sum_{j=1}^p f_j \ln \pi_j \quad (59)$$

を最大とする解を求める。ラグランジュの乗数を λ とし

$$Q = \ln L + \lambda \left(1 - \sum_{j=1}^p \pi_j \right) \quad (60)$$

を考える。これをパラメタで偏微分すると

$$\frac{\partial Q}{\partial \pi_k} = \frac{f_k}{\pi_k} - \lambda \quad (61)$$

を得るが、これを 0 と置くことにより、

$$\pi_k = \frac{1}{\lambda} f_k \quad (62)$$

となる。ここで、上式を全て加えると

$$\sum_{j=1}^p \pi_k = \sum_{j=1}^p \frac{1}{\lambda} f_k = \frac{1}{\lambda} \sum_{j=1}^p f_k \quad (63)$$

となるが、Eq.(58) より（これは Q のラグランジュの乗数に関する偏微分を 0 と置いた結果である。）これは 1 である。したがって

$$\frac{1}{\lambda} \sum_{k=1}^p f_k = 1 \quad (64)$$

より、

$$\lambda = \sum_{k=1}^p f_k \quad (65)$$

を得る。また、

$$\pi_k = \frac{f_k}{\sum_{k=1}^p f_k} \quad (66)$$

である。

2.6 例

先の例にあらわれた対数尤度

$$\ln L = \sum_{j=1}^p f_j \ln \pi_j \quad (67)$$

を $r2 \times p$ の行列 B と $r2 \times 1$ のベクトル c を用いて

$$B\pi = c \quad \text{or} \quad \sum_{j=1}^p b_{lj}\pi_j = c_l, \quad l = 1, 2, \dots, r2 \quad (68)$$

という制約条件の下で最大化することを考える。ただし、全成分の和が 1 となるという制約は上記の制約の中に含まれているものとする。(たとえば、 \mathbf{B} 行列の第 1 行が全て 1 で、 $c_1 = 1$ とする。)

Let us define the constrained log likelihood as

$$Q = \ln L - \sum_{l=1}^{r2} \lambda_l \left(\sum_{j=1}^p b_{lj} \pi_j - c_l \right), \quad (69)$$

where $\lambda_l, l = 1, 2, \dots, r2$, is the Lagrangean multipliers. By differentiating Q with respect to π_j and setting the result equal to zero, we have,

$$\pi_j = \frac{f_j}{\sum_{l=1}^{r2} \lambda_l b_{lj}}, \quad j = 1, 2, \dots, p. \quad (70)$$

Therefore, the original form B constraints can be written as

$$\sum_{j=1}^p b_{kj} \pi_j = \sum_{j=1}^p \frac{b_{kj} f_j}{\sum_{l=1}^{r2} \lambda_l b_{lj}} = c_k \quad (71)$$

or,

$$\sum_{j=1}^p \frac{b_{kj} f_j}{\sum_{l=1}^{r2} \lambda_l b_{lj}} - c_k = 0. \quad (72)$$

for $k = 1, 2, \dots, r2$. This is the system of $r2$ equations for unknown λ'_l s, and the numerical solution may be obtained by Newton's method. The necessary derivative matrix of the above equations is

$$\frac{\partial g_k}{\partial \lambda_m} = - \sum_{j=1}^p \frac{f_j b_{kj} b_{mj}}{(\sum_{l=1}^{r2} \lambda_l b_{lj})^2}, \quad (73)$$

for $k, m = 1, 2, \dots, r2$ where g_k stands for the left hand side of the k -th equation in Eq.(72). After obtaining the value of the Lagrangean multipliers, the maximum likelihood estimates of π can be calculated by substituting those values into equation Eq.(70).

Note that the same can be accomplished more easily if we use the equivalent form A constraints.

2.7 例

Fuzzy c-means Clustering

Objective function is

$$f = \sum_{i=1}^I \sum_{k=1}^K m_{ik}^q \|y_{(i)} - c_{(k)}\|^2 \quad (74)$$

$$= \sum_{i=1}^I \sum_{k=1}^K m_{ik}^q d_{ik} \quad (75)$$

where

$$d_{ik} = \|y_{(i)} - c_{(k)}\|^2 \quad (76)$$

The objective is to find those m_{ik} $i = 1, 2, \dots, I$, $k = 1, 2, \dots, K$, which minimize f above subject to

$$\sum_{k=1}^K m_{ik} = 1 \quad (77)$$

Let L_i , $i = 1, 2, \dots, I$, be the Lagrange multipliers and we minimize

$$f1 = f - \sum_{i=1}^I L_i \left(\sum_{k=1}^K m_{ik} - 1 \right) \quad (78)$$

with respect to m_{ik} s and L_i s.

Now

$$\frac{\partial f1}{\partial m_{ik}} = d_{ik} q m_{ik}^{q-1} - L_i. \quad (79)$$

By setting above equal to zero, we have

$$m_{ik}^{q-1} = \frac{L_i}{d_{ik} q} \quad (80)$$

and, accordingly,

$$m_{ik} = \left(\frac{L_i}{d_{ik} q} \right)^{1/(q-1)} \quad (81)$$

$$= \left(\frac{L_i}{q} \right)^{1/(q-1)} \times \frac{1}{d_{ik}^{1/(q-1)}} \quad (82)$$

Therefore, by Eq.(77), we have

$$\sum_{k=1}^K m_{ik} = \sum_{k=1}^I \left[\left(\frac{L_i}{q} \right)^{1/(q-1)} \times \frac{1}{d_{ik}^{1/(q-1)}} \right] \quad (83)$$

$$= \left(\frac{L_i}{q} \right)^{1/(q-1)} \times \sum_{k=1}^K \frac{1}{d_{ik}^{1/(q-1)}} \quad (84)$$

$$= 1 \quad (85)$$

and solving the above equation for $\left(\frac{L_i}{q} \right)^{1/(q-1)}$, we have

$$\left(\frac{L_i}{q} \right)^{1/(q-1)} = \frac{1}{\sum_{k=1}^K \frac{1}{d_{ik}^{1/(q-1)}}} \quad (86)$$

$$= \frac{1}{\sum_{a=1}^K \frac{1}{d_{ia}^{1/(q-1)}}} \quad (87)$$

Therefore, by substituting above in Eq.(82), we have

$$m_{ik} = \frac{1}{\sum_{a=1}^K \frac{1}{d_{ia}^{1/(q-1)}}} \times \frac{1}{d_{ik}^{1/(q-1)}} \quad (88)$$

$$= \frac{\frac{1}{d_{ik}^{1/(q-1)}}}{\sum_{a=1}^K \frac{1}{d_{ia}^{1/(q-1)}}} \quad (89)$$

$$= \frac{1}{\sum_{a=1}^K \frac{d_{ik}^{1/(q-1)}}{d_{ia}^{1/(q-1)}}} \quad (90)$$

$$= \frac{1}{\sum_{a=1}^K \left(\frac{d_{ik}}{d_{ia}}\right)^{1/(q-1)}} \quad (91)$$

2.8 例

ラグランジュの乗数が 0 となる例。

•

$$f(x, y) = x^2 + 4xy + 4y^2 \quad (92)$$

$$\text{subject to : } x^2 + 2y^2 = 4 \quad (93)$$

Solution

det=0

•

$$f(x, y) = x^2 + y^2 \quad (94)$$

$$\text{subject to : } y = x^2 \quad (95)$$

Solution

•

$$f(x, y) = x^2 + y^2 \quad (96)$$

$$\text{subject to : } (x-1)^2 + y^2 = 1 \quad (97)$$

Solution

•

$$f(x, y, z) = xyz \quad (98)$$

$$\text{subject to : } x^2 + 2y^2 + 3z^2 = 6 \quad (99)$$

There are 14 different solutions.

3 双対系を用いた解

$p \times 1$ のパラメタ β にかかる線形の制約は、以下の二つの形に書くことができる。

Form A 制約 (A, d)

$$\beta = A\gamma + d, \quad (100)$$

ただし、 A は階数が r_1^* である $p \times r_1$ の行列、 d は $p \times 1$ のベクトル、 γ は $r_1 \times 1$ の任意のベクトルである。

Form B 制約 (B, c)

$$B\beta = c, \quad (101)$$

ただし、 B は、階数が r_2^* である $r_2 \times p$ の行列、 c は $r_2 \times 1$ のベクトルであるが、この制約が意味を持つためには、 c が B の列空間の中になければならない。

この二つの線形制約の形は互いに双対系をなしており、与えられた線形制約をどちらの形でも表現することが出来る。これらの二つの制約のどちらかが与えられた場合、他の一方への変換は以下のようにして行うことができる。

Form A から Form B への変換

階数が r_1^* の form A の制約 (A, d) が与えられたときに、それと同値の階数 $p - r_1^*$ の form B の制約 (B_A, c_A) は

$$B_A = A_R, \quad c_A = A_R d, \quad (102)$$

で与えられる。ただし、 A_R は $(p - r_1^*) \times p$ の行列であり、

$$Q_A = I_p - A(A'A)^{-1}A' = A_L A_R. \quad (103)$$

を満たすものである。(一意には定まらない。) すなわち、 B_A 行列として、 A 行列の列空間に直交するような、階数が r_1^* のフルランクの行列を用いればよい。

したがって

$$B_A \beta = A_R(A\gamma + d) \quad (104)$$

$$= A_R A \gamma + A_R d = \mathbf{0} + c_A \quad (105)$$

となる。

Form B から Form A への変換

階数が r_2^* の Form B の制約 (B, c) が与えられた場合、それと同値の階数 $p - r_2^*$ の form A の制約 (A_B, d_B) は

$$A_B = B_L, \quad d_B = (B'B)^{-1}B'c, \quad (106)$$

となる。ただし、 B_L は $p \times p - r_2^*$ の行列で

$$Q_{B'} = I_p - B'(BB')^{-1}B = B_L B_R. \quad (107)$$

を満たすものである。(一意には定まらない。) すなわち、 A_B 行列として、 B 行列の行空間に直交するような、階数が $p - r_2^*$ のフルランクの行列を用いればよい。したがって

$$B\beta = B(A_B\gamma + d_B) = B(B_L\gamma + (B'B)^{-1}B'c) \quad (108)$$

$$= BB_L\gamma + B(B'B)^{-1}B'c = \mathbf{0} + c \quad (109)$$

となる。(整合性のある制約の場合)

一般に、Form B の形の制約は、その意味することを理解しやすい形であるが、前節に示したように、ラグランジュの方法を用いたその取り扱いが面倒である。片や Form A は、その意味するところの理解は難しいが、解がきわめて簡単となり、特に線形モデルの場合は、求める $\hat{\beta}$ は

$$\hat{\beta} = (A'X'XA)^{-1}A'X'y + d \quad (110)$$

となる。

先の Eq.(6) 示した (L, c) に対応する (A, d) は以下のようになる。

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad d = \mathbf{0} \quad (111)$$