

Fisher Information Matrix

ryamada

2017年2月25日

- 1 はじめに Introduction
- 2 フィッシャー情報量のいろいろな見方 Multiple ways to interpret Fisher information
 - 2.1 分布の球面上配置 半径 2 の球面 Distributions on the surface of sphere with radius 2.
 - 2.2 対数確率密度関数・対数尤度関数の偏微分の分散・共分散行列(ばらつきの期待値) Variance/Covariance matrix (expected of variability) of log probability density function/log likelihood function
 - 2.3 対数尤度関数の二階の偏微分の期待値(が作る正方行列) Expected of 2nd order partial differentiation of log-likelihood function
 - 2.4 リーマン計量 Riemannian metrics
 - 2.5 球面上の 2 つの偏微分の積でもあり、異なる 2 つの変換関数の偏微分の積でもある Product of two partial derivative. Also it is product of a partial derivative of one function and a partial derivative of another function.
- 3 具体例 三項分布 Example, three-nomial.
 - 3.1 格子 Lattice
 - 3.2 フィッシャー情報量の成分と格子 Elements of Fisher information matrix and lattice
- 4 参考 導出 Appendix
 - 4.1 α

1 はじめに Introduction

- 分布を点として扱うと、分布たちが空間に並んでいる様子を扱うことになる。We treat distribution as points in a space and evaluate how distributions are arranged in the space.
- 並んでいる分布を捕まえるには、分布に座標を与えるとよい Coordinates for distributions are convenient to understand the arrangement.
- 座標を与えると、座標の変化に応じて分布間の違いがどのように増減するかが気になる Once coordinates are given, changes of dissimilarity among distributions along the coordinates are of interest.
- その座標に沿った分布の違いの増減を表すのがフィッシャー情報量である Fisher information describes the changes of measure along the coordinates.
- フィッシャー情報量の表現式は色々に変えることができ、その式を読むことで、フィッシャー情報量の異なる意味が読み取れる Fisher information can be expressed in multiple ways of formula, each of which indicates what Fisher information is.

2 フィッシャー情報量のいろいろな見方 Multiple ways to interpret Fisher information

- 分布を球面上に配置し、その局所での微小ベクトルの長さの 2 乗がフィッシャー情報量であることを説明したのち、それを式変形することで、別の式表現とし、その意味を説明することで、フィッシャー情報量のいろいろな見方を示す。The followings include; distributions can be placed on a surface of sphere; the square of infinitesimal vector's length and/or inner product of two those vectors are the elements of Fisher information matrix; the formula of Fisher information are changed; they are helpful to understand what Fisher information stands for.

- 式変形のやり方は示さないが、以下を使うことで式変形することができる。In this textbook, formula transformations are not fully shown but the majority of transformations uses the equations below.

\$\$ = \backslash = \backslash

\$\$

2.1 分布の球面上配置 半径 2 の球面 Distributions on the surface of sphere with radius 2.

- 分布 $P = \{p_1, \dots, p_k\}; k = 1, \dots, \infty$ は球面上の点 $Q = \{q_i = \sqrt{p_i}\}; \sum_i q_i^2 = 1$ とみなせる
 $Q = \{q_i = \sqrt{p_i}\}; \sum_i q_i^2 = 1$ describes that any distribution $P = \{p_1, \dots, p_k\}; k = 1, \dots, \infty$ can be placed on a sphere.
- 分布のパラメタ t がわずかに変化したときの、局所球面上の微小距離 dl は The small distance on the sphere surface dl with small change of paramter t , dt is;

$$\frac{dl}{d\theta} \frac{dl}{d\theta} = \int \frac{d\sqrt{p}}{d\theta} \frac{d\sqrt{p}}{d\theta} dx$$

- フィッシャー情報量はこの 4 倍。分布に対応する点が、単位球面にあるのではなくて、半径 2 の球面にあることに相当させて、以下ようになる。This is the square of length on a unit sphere and four times this expression is Fisher information, therefore Fisher information is the corresponding on the sphere with radius 2.

$$FisherInfo = 4 \frac{dl}{d\theta} \frac{dl}{d\theta} = \int \frac{d2\sqrt{p}}{d\theta} \frac{d2\sqrt{p}}{d\theta} dx$$

- 実際、分布が k 個の変数で表された分布のフィッシャー情報量は、 $k \times k$ 行列である Fisher information is $k \times k$ matrix when p is k -variate function.

$$FisherInfo_{i,j} = 4 \frac{dl}{d\theta_i} \frac{dl}{d\theta_j} = \sum_w \frac{d2\sqrt{p(w)}}{d\theta_i} \frac{d2\sqrt{p(w)}}{d\theta_j} = \int \frac{d2\sqrt{p}}{d\theta_i} \frac{d2\sqrt{p}}{d\theta_j} dx$$

- $\sum p(w) = 1$ もしくは $\int p(x) dx = 1$ を $2\sqrt{p}$ に変換したときの、局所の微分微小ベクトルの『ピタゴラスの定理的な和』。 $\sum_w \int dx$ はすべての $p(w), p(x)$ について足し合わせることで『ピタゴラスの定理』で斜辺の長さの二乗を出す操作 The summation \sum over the number of elements p_i means the square of oblique segments as the Pythagorean theorem indicates, after the transformation $2\sqrt{p}$. \int is the version of the Pythagorean theorem calculation for infinite dimension.

2.2 対数確率密度関数・対数尤度関数の偏微分の分散・共分散行列(ばらつきの期待値) Variance/Covariance matrix (expected of variability) of log probability density function/log likelihood function

- 式変形をすると、以下のようにも表せる。最後の p が「期待値」であることを示している。Fisher information can be expressed as below. The last p indicates “expected value”.

$$FisherInfo_{i,j} = \int \frac{\partial \log p}{\partial \theta_i} \frac{\partial \log p}{\partial \theta_j} p dx$$

- $i = j$ のとき、上式は $(\frac{\partial \log p}{\partial \theta_i})^2$ (対数確率密度関数・対数尤度関数の偏微分の二乗)を計算しており、それに p の重みをつけて平均を取っている。実は、対数確率密度関数の偏微分の平均は0なので、これは、分散(の期待値)になっている。When $i = j$ the formula above calculates 2nd moment around 0. Actually the mean of $\frac{\partial \log p}{\partial \theta_i}$ is zero, therefore, the 2nd moment is variance.
- $i \neq j$ のときは、共分散 When $i \neq j$ it is covariance.
- ある確率分布からサンプルが得られているとき、真実の確率分布のもとで対数尤度を計算すると、その平均値は0となり、分散共分散はフィッシャー情報量になることを意味する。This means, when we have samples from a distribution and calculate log likelihood for the TRUE distribution, then its average should be 0 and its variance/covariance matrix is Fisher information matrix.
- フィッシャー情報量が小さいならば、サンプルから推定される確率分布のパラメタの精度がよく、大きいならば、精度が悪いことを意味する。When Elements of Fisher information is small, samples will indicate the true distribution with precision and vise versa.

2.3 対数尤度関数の二階の偏微分の期待値(が作る正方形行列) Expected of 2nd order partial differentiation of log-likelihood function

$$FisherInfo_{i,j} = \int \frac{\partial \log p}{\partial \theta_i} \frac{\partial \log p}{\partial \theta_j} p dx = \int (\frac{\partial}{\partial \theta_i} \frac{\partial}{\partial \theta_j} \log p) p dx$$

- 式が示すとおり、対数尤度関数の2階の微分の確率重み付き平均(期待値)となっている。二階の微分は、伸び縮み具合が変化しているかを表す量であり、対数尤度関数の値を高さとして描いたなら、曲率に相当する量である。The formula indicate expected of 2nd order partial differentiation. The 2nd order differentiation stands for how changes the 1st order differentiation, meaning how quickly elongation factor changes. When the function is drawn as "high", then the 2nd order differentiation describes curvature.

2.4 リーマン計量 Riemannian metrics

$$FisherInfo_{i,j} = \int \frac{\partial \log p}{\partial \theta_i} \frac{\partial \log p}{\partial \theta_j} p dx = \int (\frac{\partial}{\partial \theta_i} \frac{\partial}{\partial \theta_j} \log p) p dx$$

- フィッシャー計量は、パラメタで表現された関数について、パラメタのペアごとに値を持った、対称正方形行列である。Fisher information is a symmetric square matrix whose elements correspond to parameter pairs.

- k 次元多様体は、各所に $k \times k$ なる計量行列を持ったもの(局所において、微小ベクトルの内積を定める)であり、その計量行列をリーマン計量と言うが、フィッシャー情報量は、このリーマン計量が満足する性質を持っているので、フィッシャー情報量は、確率密度分布を多様体とみなしたときのリーマン計量である、と言う。k-dimensional manifold has $k \times k$ measure matrix that defines inner product of small vectors (tangent vectors), which is Riemannian measure. Fisher information matrix satisfies the features and it is Riemannian measure for manifold of distributions.

2.5 球面上の2つの偏微分の積でもあり、異なる2つの変換関数の偏微分の積でもある Product of two partial derivative. Also it is product of a partial derivative of one function and a partial derivative of another function.

$$FisherInfo_{i,j} = \int \frac{\partial \log p}{\partial \theta_i} \frac{\partial \log p}{\partial \theta_j} p dx = \int \frac{\partial \log p}{\partial \theta_i} \frac{\partial p}{\partial \theta_j} dx = \int \frac{\partial 2\sqrt{p}}{\partial \theta_i} \frac{\partial 2\sqrt{p}}{\partial \theta_j} dx$$

- 上段は、対数尤度関数の偏微分を2つ掛け合わせたものに、さらにpをかけて積分しているので、偏微分の掛け合わせの「期待値」を表す。The top line: Product of two partial derivatives of log likelihood function that is multiplied by p, that is integrated. It means the expected of the product of partial derivatives of log likelihood function.
- 他方、中段と下段は、2つの関数の偏微分の積を積分しているから、「期待値」ではない The middle and bottom lines don't have p and it means they are not "the expected".
- 下段は、 $2\sqrt{p}$ という関数の偏微分同士の積を、すべてのxについて足し合わせており、「ピタゴラスの定理的な」「長さの二乗」を表していた。具体的には半径2の球面上の微小距離の2乗である。The bottom is the sum of the square of $2\sqrt{p}$ over all axes of space, which means the Pythagorean summation, i.e., the square of oblique segment. The calculation is performed on the surface of sphere of radius 2.
- 中段は「期待値」でもないが、いわゆるピタゴラスの定理的な長さの二乗というわけでもない。あえて言えば、ピタゴラスの定理的な長さの二乗は、相互に直交する長さの二乗の和に分解できるが、その分解成分を「二乗～正方形」ではなく、「長方形」として計算して、結果として、同じ値が返るようになったものである The middle indicates that it is not "the expected" of something. It is not the squared distance in the Pythagorean way, either. If we consider the square of something as the area of square, the middle calculation is the product of two different length, meaning the area of rectangle as below.
- $\frac{\partial 2\sqrt{p}}{\partial \theta} \times \frac{\partial 2\sqrt{p}}{\partial \theta}$ なる正方形と同じ面積を持つ $\frac{\partial \log p}{\partial \theta_i} \times \frac{\partial p}{\partial \theta_j}$ なる長方形の面積で表現しなおしたもの、と言える。
 $\frac{\partial 2\sqrt{p}}{\partial \theta} \times \frac{\partial 2\sqrt{p}}{\partial \theta}$ and $\frac{\partial \log p}{\partial \theta_i} \times \frac{\partial p}{\partial \theta_j}$ are the same but the former is the area of square and the latter is the area of rectangle.
- さらに一般化すると It is generalized further as;

$$FisherInfo_{i,j} = \int \frac{\partial \log p}{\partial \theta_i} \frac{\partial \log p}{\partial \theta_j} p dx = \int \frac{\partial \log p}{\partial \theta_i} \frac{\partial p}{\partial \theta_j} dx = \int \frac{\partial 2\sqrt{p}}{\partial \theta_i} \frac{\partial 2\sqrt{p}}{\partial \theta_j} dx = \int \frac{\partial l^{(\alpha)}}{\partial \theta_i} \frac{\partial l^{(-\alpha)}}{\partial \theta_j} dx$$

ただし where,

$$l^{(\alpha)} = \frac{2}{1-\alpha} p^{\frac{1-\alpha}{2}}; \alpha \neq 1 = \log p; \alpha = 1$$

- このことから、 $2\sqrt{p}$ は $\alpha = 0$ 、 p は $\alpha = -1$ 、 $\log p$ は $\alpha = 1$ の表現と呼ばれる。 $\alpha = 0, \pm 1$ のみを対象にするので、気にする必要はほとんどないが、情報幾何に関する文書に登場する記法なので、参考のために挙げておく。This expression says that $2\sqrt{p}$ is the form of $\alpha = 0$ and p is the form of $\alpha = -1$. We don't have to worry about this α forms much because we only care $\alpha = 0, \pm 1$, but this generalization might be helpful to understand duality that is the topic of the next section.

3 具体例 三項分布 Example, three-nomial.

$$P = \{p_1, p_2, p_3\}$$

3.1 格子 Lattice

- 3通りのパラメタの取り方をする Let's take three parameterization ways.
- 1つ目 1st

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ 1 - \alpha_1 - \alpha_2 \end{pmatrix}$$

- 2つ目 2nd

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} \frac{e^{\beta_1}}{1 + e^{\beta_1} + e^{\beta_2}} \\ \frac{e^{\beta_2}}{1 + e^{\beta_1} + e^{\beta_2}} \\ \frac{1}{1 + e^{\beta_1} + e^{\beta_2}} \end{pmatrix}$$

書き換えると、they are:

$$\begin{pmatrix} \log p_1 \\ \log p_2 \\ \log p_3 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ 0 \end{pmatrix} - \log(1 + e^{\beta_1} + e^{\beta_2}) \begin{pmatrix} \log \frac{p_1}{p_3} \\ \log \frac{p_2}{p_3} \\ 1 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

- 3つ目 3rd

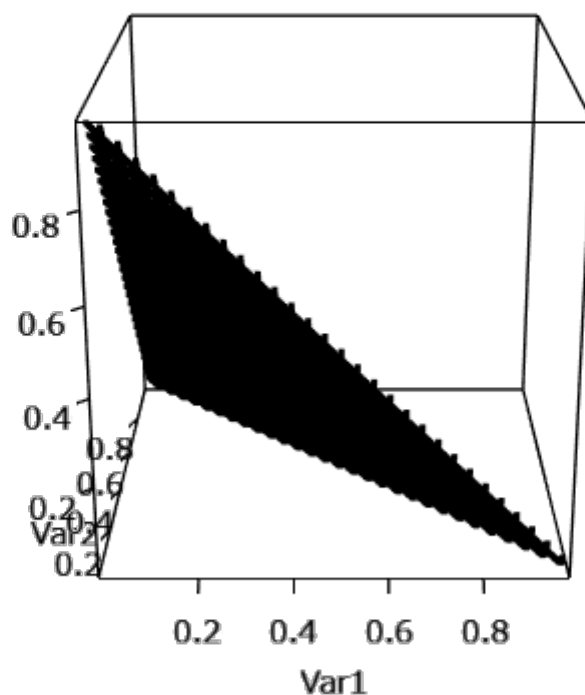
$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} (\frac{\gamma_1}{2})^2 \\ (\frac{\gamma_2}{2})^2 \\ 1 - ((\frac{\gamma_1}{2})^2 + (\frac{\gamma_2}{2})^2) \end{pmatrix}$$

書き換えると、they are:

$$\begin{pmatrix} 2\sqrt{p_1} \\ 2\sqrt{p_2} \\ 2\sqrt{p_3} \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \sqrt{4 - \gamma_1^2 - \gamma_2^2} \end{pmatrix}$$

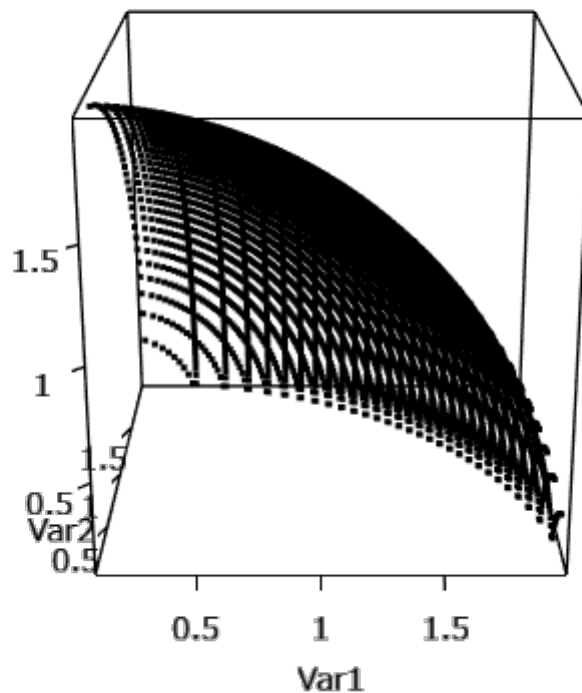
- 図示してみる Draw them.
- $p_1 + p_2 + p_3 = 1$ の平面に (α_1, α_2) の格子を描く Plane of $p_1 + p_2 + p_3 = 1$ with coordinate lattice of (α_1, α_2) .

```
tmp.p1 <- tmp.p2 <- seq(from=0, to=1, length=31)
tmp.p1. <- tmp.p2. <- seq(from=0, to=1, length=301)
tmp.p1.p2A <- as.matrix(expand.grid(tmp.p1, tmp.p2.))
tmp.p1.p2B <- as.matrix(expand.grid(tmp.p1., tmp.p2))
tmp.p1.p2 <- rbind(tmp.p1.p2A, tmp.p1.p2B)
tmp.p1.p2.p3 <- cbind(tmp.p1.p2, 1-apply(tmp.p1.p2, 1, sum))
ok <- which(apply(tmp.p1.p2.p3, 1, min)>0)
P <- tmp.p1.p2.p3[ok,]
library(rgl)
plot3d(P)
```



- $p_1^2 + p_2^2 + p_3^2 = 4$ の球面に (α_1, α_2) の格子を描く Sphere surface of $p_1^2 + p_2^2 + p_3^2 = 4$ with coordinate lattice of (α_1, α_2) .

```
Q <- 2*sqrt(tmp.p1.p2.p3[ok,])
plot3d(Q)
```



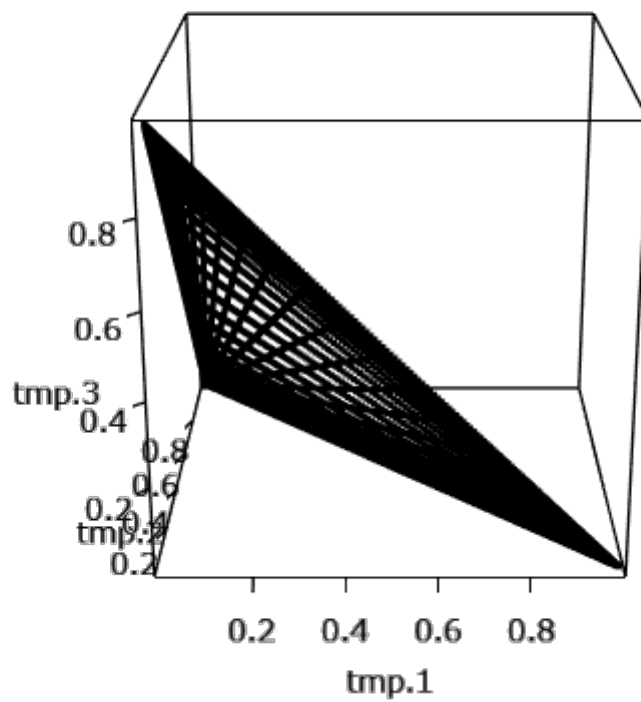
- $p_1 + p_2 + p_3 = 1$ の平面に (β_1, β_2) の格子を描く Plane of $p_1 + p_2 + p_3 = 1$ with coordinate lattice of (β_1, β_2) .

```

tmp.p1 <- tmp.p2 <- seq(from=-5, to=5, length=31)
tmp.p1. <- tmp.p2. <- seq(from=-5, to=5, length=301)
tmp.p1.p2A <- as.matrix(expand.grid(tmp.p1, tmp.p2.))
tmp.p1.p2B <- as.matrix(expand.grid(tmp.p1., tmp.p2))
tmp.p1.p2 <- rbind(tmp.p1.p2A, tmp.p1.p2B)

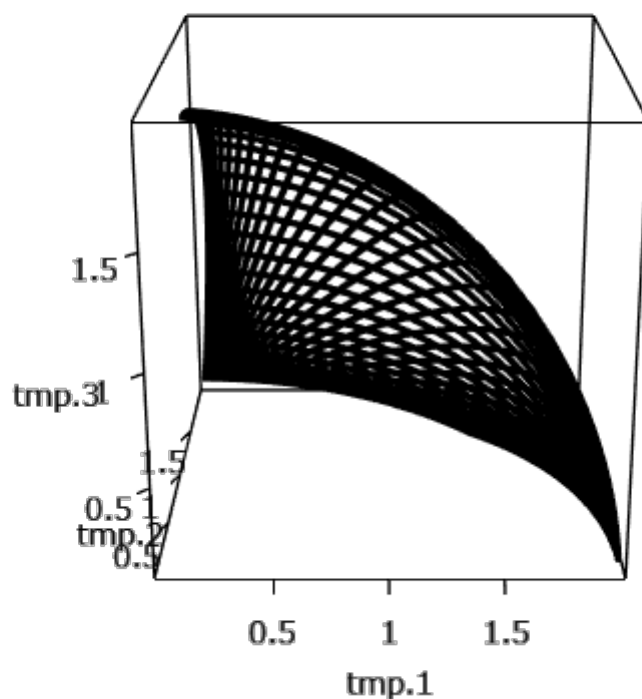
tmp.3 <- 1/(1+exp(tmp.p1.p2[, 1]) + exp(tmp.p1.p2[, 2]))
tmp.1 <- exp(tmp.p1.p2[, 1])*tmp.3
tmp.2 <- exp(tmp.p1.p2[, 2])*tmp.3
tmp.p1.p2.p3 <- cbind(tmp.1, tmp.2, tmp.3)
ok <- which(apply(tmp.p1.p2.p3, 1, min)>0)
P <- tmp.p1.p2.p3[ok, ]
library(rgl)
plot3d(P)

```



- $p_1^2 + p_2^2 + p_3^2 = 4$ の球面に (β_1, β_2) の格子を描く Sphere surface of $p_1^2 + p_2^2 + p_3^2 = 4$ with coordinate lattice of (β_1, β_2) .

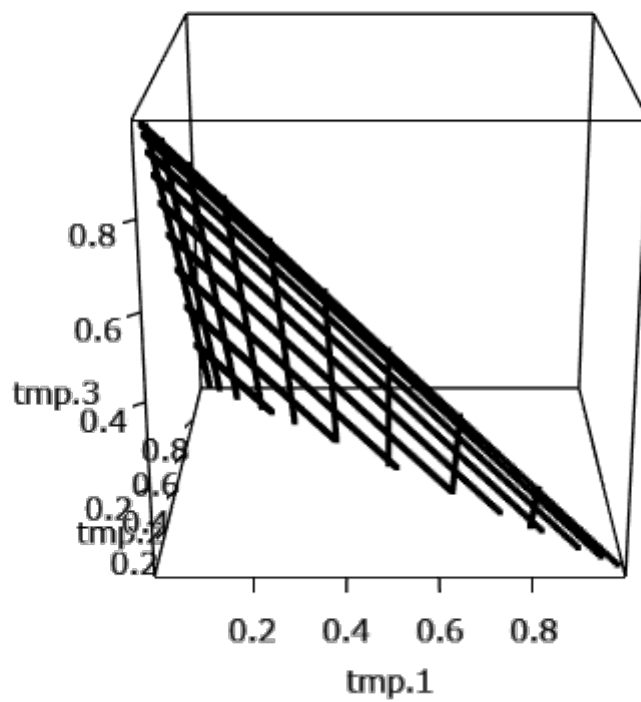
```
Q <- 2*sqrt(tmp.p1.p2.p3[ok,])
plot3d(Q)
```

- $p_1 + p_2 + p_3 = 1$ の平面に (γ_1, γ_2) の格子を描く Plane of $p_1 + p_2 + p_3 = 1$ with coordinate lattice of (γ_1, γ_2) .

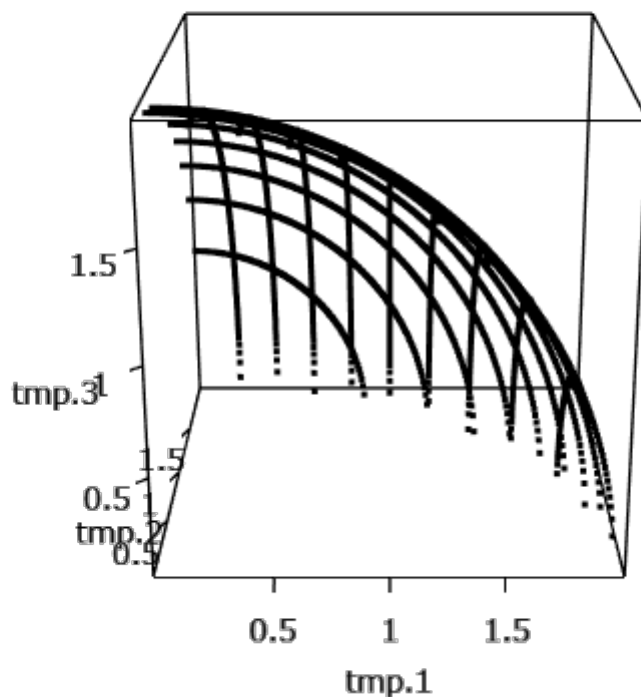
```
tmp.p1 <- tmp.p2 <- seq(from=0, to=2, length=11)
tmp.p1. <- tmp.p2. <- seq(from=0, to=2, length=301)
tmp.p1.p2A <- as.matrix(expand.grid(tmp.p1, tmp.p2.))
tmp.p1.p2B <- as.matrix(expand.grid(tmp.p1., tmp.p2))
tmp.p1.p2 <- rbind(tmp.p1.p2A, tmp.p1.p2B)

tmp.1 <- tmp.p1.p2[, 1]^2/4
tmp.2 <- tmp.p1.p2[, 2]^2/4
tmp.3 <- 1- tmp.1-tmp.2
tmp.p1.p2.p3 <- cbind(tmp.1, tmp.2, tmp.3)
ok <- which(apply(tmp.p1.p2.p3, 1, min)>0)
P <- tmp.p1.p2.p3[ok, ]
library(rgl)
plot3d(P)
```



- $p_1^2 + p_2^2 + p_3^2 = 4$ の球面に (γ_1, γ_2) の格子を描く Sphere surface of $p_1^2 + p_2^2 + p_3^2 = 4$ with coordinate lattice of (γ_1, γ_2) .

```
Q <- 2*sqrt(tmp.p1.p2.p3[ok,])
plot3d(Q)
```



3.2 フィッシャー情報量の成分と格子 Elements of Fisher information matrix and lattice

- フィッシャー情報量は以下の式で表される Elements of Fisher information matrix

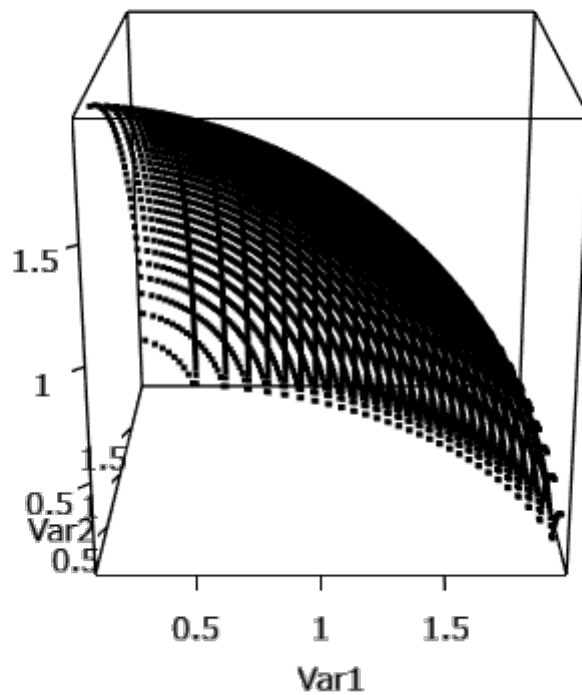
$$g(ij) = \sum_{k=1}^3 \frac{\partial^2 \sqrt{p}}{\partial \theta_i} \frac{\partial^2 \sqrt{p}}{\partial \theta_j}$$

- この意味するところは、次のようなことである。This means:
- フィッシャー情報行列の対角成分 $g(ii)$ は、半径 2 の球面に (θ_1, θ_2) で座標を取ったときに(そのような格子を描いたときに)、 θ_i 方向の格子のマス目の幅の長短の二乗相当量のことである。The diagonal elements of Fisher information matrix is something similar to the square of length of segments of lattice of (θ_1, θ_2) that are drawn on the sphere surface of radius 2.
- また、 $g(ij); i \neq j$ は、 θ_i 方向と θ_j 方向とが作る角に関係する値である。 $\cos \phi = \frac{g(ij)}{\sqrt{g(ii)g(jj)}}$ である。 $g(ij); i \neq j$ is a value related with the angle of two coordinate direction vectors; $\cos \phi = \frac{g(ij)}{\sqrt{g(ii)g(jj)}}$.
- 実際、 α について、 $G = (g(ij))$ を求めると以下のようにになる。 $G = (g(ij))$ for α is;

$$G_{\alpha} = (g_{\alpha}(ij)) = \begin{pmatrix} \frac{1}{\alpha_1} + \frac{1}{(1-\alpha_1-\alpha_2)}, & \frac{1}{(1-\alpha_1-\alpha_2)} \\ \frac{1}{(1-\alpha_1-\alpha_2)}, & \frac{1}{\alpha_2} + \frac{1}{(1-\alpha_1-\alpha_2)} \end{pmatrix} = \begin{pmatrix} \frac{1}{p_1} + \frac{1}{p_3}, & \frac{1}{p_3} \\ \frac{1}{p_3}, & \frac{1}{p_2} + \frac{1}{p_3} \end{pmatrix}$$

- 以下の図と比べてみる Read the description with the 3d plot below.
 - フィッシャー情報行列の対角成分 $g(ii)$ は、半径 2 の球面に (θ_1, θ_2) で座標を取ったときに(そのような格子を描いたときに)、 θ_i 方向の格子のマス目の幅の長短の二乗相当量のことである。The diagonal elements $g(ii)$ are the values corresponding to the square of segments of lattice of (θ_1, θ_2)
 - $g(11)$ 成分は $\frac{1}{p_1} + \frac{1}{p_3}$ 。 The element $g(11)$ is $\frac{1}{p_1} + \frac{1}{p_3}$.
 - これは、 p_2 が等しいとき、 p_1, p_3 が等しいときに値が最小になり、 p_1, p_3 の値が異なると大きくなる_ことを意味する。 $p_2 = 0$ の円弧に沿ってマス目の大小変化を見ると、確かに $p_1 = p_3 = 0.5, p_2 = 0$ で最も密になっている See the line where p_2 is fixed, for example $p_2 = 0$, then the segment is shortest when $p_1 = p_3$. The segments get longer with the difference between p_1 and p_3 get bigger.
 - また、 $g(ij); i \neq j$ は、 θ_i 方向と θ_j 方向とが作る角のコサイン相当の値である。 $g(ij); i \neq j$ is related with cosine of angles between two directions, direction θ_1 and direction θ_j .
 - 非対角成分 $g(12)$ は $\frac{1}{p_3}$ である。確かに p_3 が 0 に近いときには、格子の成す角が狭まり、 p_3 が 1 に近いときには、格子は直角に近い。 $g(12) = \frac{1}{p_3}$. When p_3 is close to zero, the angle is small and when p_3 is close to 1, then the angle is close to $\pi/2$.
 - $\$ = \frac{1}{p_3} \frac{1}{\sqrt{(\frac{1}{p_1} + \frac{1}{p_3})(\frac{1}{p_2} + \frac{1}{p_3})}} = \sqrt{\frac{p_1 p_2}{(p_1 + p_3)(p_2 + p_3)}}$ なので、確かに、 $p_3 = 1$ のとき、この値は 0 となり
- $p_3 = 0, p_1 = p_2 = 0.5$ のとき、この値は 1 になる。 When $p_3 = 1$ the value $\sqrt{\frac{p_1 p_2}{(p_1 + p_3)(p_2 + p_3)}}$ is 0 and when
- $p_3 = 0, p_1 = p_2 = 0.5$, the values is 1.

```
tmp.p1 <- tmp.p2 <- seq(from=0, to=1, length=31)
tmp.p1. <- tmp.p2. <- seq(from=0, to=1, length=301)
tmp.p1.p2A <- as.matrix(expand.grid(tmp.p1, tmp.p2.))
tmp.p1.p2B <- as.matrix(expand.grid(tmp.p1., tmp.p2))
tmp.p1.p2 <- rbind(tmp.p1.p2A, tmp.p1.p2B)
tmp.p1.p2.p3 <- cbind(tmp.p1.p2, 1-apply(tmp.p1.p2, 1, sum))
ok <- which(apply(tmp.p1.p2.p3, 1, min)>0)
P <- tmp.p1.p2.p3[ok, ]
Q <- 2*sqrt(tmp.p1.p2.p3[ok, ])
plot3d(Q)
```



4 参考 導出 Appendix

4.1 α

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ 1 - \alpha_1 - \alpha_2 \end{pmatrix}$$

$$\begin{pmatrix} 2\sqrt{p_1} \\ 2\sqrt{p_2} \\ 2\sqrt{p_3} \end{pmatrix} = \begin{pmatrix} 2\sqrt{\alpha_1} \\ 2\sqrt{\alpha_2} \\ 2\sqrt{1 - \alpha_1 - \alpha_2} \end{pmatrix}$$

$$\frac{\partial}{\partial \alpha_1} \begin{pmatrix} 2\sqrt{p_1} \\ 2\sqrt{p_2} \\ 2\sqrt{p_3} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{\alpha_1}} \\ 0 \\ \frac{-1}{\sqrt{1 - \alpha_1 - \alpha_2}} \end{pmatrix} \frac{\partial}{\partial \alpha_2} \begin{pmatrix} 2\sqrt{p_1} \\ 2\sqrt{p_2} \\ 2\sqrt{p_3} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{\alpha_2}} \\ \frac{-1}{\sqrt{1 - \alpha_1 - \alpha_2}} \end{pmatrix}$$

$$G_{\alpha} = (g_{\alpha}(ij)) = \begin{pmatrix} \frac{1}{\alpha_1} + \frac{1}{(1-\alpha_1-\alpha_2)}, & \frac{1}{(1-\alpha_1-\alpha_2)} \\ \frac{1}{(1-\alpha_1-\alpha_2)}, & \frac{1}{\alpha_2} + \frac{1}{(1-\alpha_1-\alpha_2)} \end{pmatrix}$$

β

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} \frac{e^{\beta_1}}{1+e^{\beta_1}+e^{\beta_2}} \\ \frac{e^{\beta_2}}{1+e^{\beta_1}+e^{\beta_2}} \\ \frac{1}{1+e^{\beta_1}+e^{\beta_2}} \end{pmatrix}$$

$$\begin{pmatrix} 2\sqrt{p_1} \\ 2\sqrt{p_2} \\ 2\sqrt{p_3} \end{pmatrix} = \begin{pmatrix} 2\sqrt{\alpha_1} \\ 2\sqrt{\alpha_2} \\ 2\sqrt{1-\alpha_1-\alpha_2} \end{pmatrix}$$

$$\frac{\partial}{\partial \alpha_1} \begin{pmatrix} 2\sqrt{p_1} \\ 2\sqrt{p_2} \\ 2\sqrt{p_3} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{\alpha_1}} \\ 0 \\ \frac{-1}{\sqrt{1-\alpha_1-\alpha_2}} \end{pmatrix} \frac{\partial}{\partial \alpha_2} \begin{pmatrix} 2\sqrt{p_1} \\ 2\sqrt{p_2} \\ 2\sqrt{p_3} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{\alpha_2}} \\ \frac{-1}{\sqrt{1-\alpha_1-\alpha_2}} \end{pmatrix}$$

$$G_{\alpha} = (g_{\alpha}(ij)) = \begin{pmatrix} \frac{1}{\alpha_1} + \frac{1}{(1-\alpha_1-\alpha_2)}, & \frac{1}{(1-\alpha_1-\alpha_2)} \\ \frac{1}{(1-\alpha_1-\alpha_2)}, & \frac{1}{\alpha_2} + \frac{1}{(1-\alpha_1-\alpha_2)} \end{pmatrix}$$