定理 21.1　非負定値対称行列 A の (非負の) 平方根，すなわち $\sqrt{A}^2 = A$ を満たす非負定値対称行列 $\sqrt{A}$ は一意に存在する．

(証明) A を n 次非負定値対称行列とすると，直交行列 $(\mathbf{e}_1\ \mathbf{e}_2 \ldots, \mathbf{e}_n)$ により次のように対角化できる (証明は後述)：

$$A = (\mathbf{e}_1\ \mathbf{e}_2 \ldots, \mathbf{e}_n) \begin{pmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{pmatrix} {}^t(\mathbf{e}_1\ \mathbf{e}_2 \ldots, \mathbf{e}_n)$$

ここで，$A\mathbf{e}_i = \alpha_i \mathbf{e}_i$，$\alpha_i \geq 0$ $(i = 1, 2, \ldots, n)$ であり，$\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ は $\mathbf{R}^n$ の正規直交基底である．そこで

$$\sqrt{A} = (\mathbf{e}_1\ \mathbf{e}_2 \ldots, \mathbf{e}_n) \begin{pmatrix} \sqrt{\alpha_1} & & & \\ & \sqrt{\alpha_2} & & \\ & & \ddots & \\ & & & \sqrt{\alpha_n} \end{pmatrix} {}^t(\mathbf{e}_1\ \mathbf{e}_2 \ldots, \mathbf{e}_n)$$

とおけば，$\sqrt{A}$ も非負定値対称行列で $\sqrt{A}^2 = A$ を満たす．よって (非負の) 平方根の存在が言えた．このとき

$$\sqrt{A}\mathbf{e}_i = \sqrt{\alpha_i}\mathbf{e}_i \ (i = 1, 2, \ldots, n)$$

が成り立っていること，すなわち，各 $\sqrt{\alpha_i}$ は $\sqrt{A}$ の固有値であり，$\mathbf{e}_i$ は固有値 $\sqrt{\alpha_i}$ に属する固有ベクトルであることに注意せよ．

一意性を示そう．非負定値対称行列 B が $B^2 = A$ を満たすとすると，任意の非負の実数 λ に対して $A - \lambda E = (B + \sqrt{\lambda}E)(B - \sqrt{\lambda}E)$ が成り立つから，$A\mathbf{e}_i = \alpha_i \mathbf{e}_i$ すなわち $(A - \alpha_i E)\mathbf{e}_i = \mathbf{0}$ より

$$(B + \sqrt{\alpha_i}E)(B - \sqrt{\alpha_i}E)\mathbf{e}_i = \mathbf{0} \quad (i = 1, 2, \ldots, n)$$

が成り立つ．ここで，もし $\alpha_i > 0$ ならば，B の非負定値性より $-\sqrt{\alpha_i}$ は B の固有値ではないから

$$(B - \sqrt{\alpha_i}E)\mathbf{e}_i = \mathbf{0} \text{ すなわち } B\mathbf{e}_i = \sqrt{\alpha_i}\mathbf{e}_i$$

を得る **(注 1)**．もし $\alpha_i = 0$ ならば $A\mathbf{e}_i = \mathbf{0}$ であるが，このとき ${}^tB = B$ より

$$\|B\mathbf{e}_i\|^2 = B\mathbf{e}_i \cdot B\mathbf{e}_i = \mathbf{e}_i \cdot B^2\mathbf{e}_i = \mathbf{e}_i \cdot A\mathbf{e}_i = \mathbf{e}_i \cdot \mathbf{0} = 0$$

よって $B\mathbf{e}_i = \mathbf{0}$ となる．以上より

$$B\mathbf{e}_i = \sqrt{\alpha_i}\mathbf{e}_i \ (i = 1, 2, \ldots, n)$$

が成り立つことになり，これは B が上記の $\sqrt{A}$ と一致することを意味する **(注 2)**．

....

(注 1)　$(B + \sqrt{\alpha_i}E)(B - \sqrt{\alpha_i}E)\mathbf{e}_i = \mathbf{0} \iff B(B - \sqrt{\alpha_i})\mathbf{e}_i = -\sqrt{\alpha_i}(B - \sqrt{\alpha_i})\mathbf{e}_i$
より，もし $(B - \sqrt{\alpha_i})\mathbf{e}_i \neq \mathbf{0}$ ならば，$-\sqrt{\alpha_i}$ が B の固有値で，$(B - \sqrt{\alpha_i})\mathbf{e}_i$ が $-\sqrt{\alpha_i}$ に属する固有ベクトルということになってしまう．

(注 2)　$\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ は $\mathbf{R}^n$ の基底であるから，$\sqrt{A}\mathbf{e}_i = B\mathbf{e}_i$ $(i = 1, 2, \ldots, n)$ ならば $\sqrt{A} = B$ である．

定理 21.2　n 次対称行列 A について，次が成り立つ．

- A が非負定値 $\iff$ 任意の $\mathbf{x} \in \mathbf{R}^n$ に対して $A\mathbf{x} \cdot \mathbf{x} \geq 0$ が成り立つ
- A が正定値 $\iff$ $\mathbf{0}$ でない任意の $\mathbf{x} \in \mathbf{R}^n$ に対して $A\mathbf{x} \cdot \mathbf{x} > 0$ が成り立つ
- A が非正定値 $\iff$ 任意の $\mathbf{x} \in \mathbf{R}^n$ に対して $A\mathbf{x} \cdot \mathbf{x} \leq 0$ が成り立つ
- A が負定値 $\iff$ $\mathbf{0}$ でない任意の $\mathbf{x} \in \mathbf{R}^n$ に対して $A\mathbf{x} \cdot \mathbf{x} < 0$ が成り立つ

(証明) A を n 次正定値対称行列とし，その固有値を $\alpha_1, \alpha_2, \ldots, \alpha_n$，それぞれに属する固有ベクトルを $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ とする．ここで，$\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ は $\mathbf{R}^n$ の正規直交基底となるようにとることができる (証明は後述)．このとき

$$\begin{aligned}
A\mathbf{x} \cdot \mathbf{x} &= A(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots x_n\mathbf{e}_n) \cdot (x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots x_n\mathbf{e}_n) \\
&= (x_1A\mathbf{e}_1 + x_2A\mathbf{e}_2 + \cdots x_nA\mathbf{e}_n) \cdot (x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots x_n\mathbf{e}_n) \\
&= (x_1\alpha_1\mathbf{e}_1 + x_2\alpha_2\mathbf{e}_2 + \cdots x_n\alpha_n\mathbf{e}_n) \cdot (x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots x_n\mathbf{e}_n) \\
&= \alpha_1{x_1}^2 + \alpha_2{x_2}^2 + \cdots + \alpha_n{x_n}^2
\end{aligned}$$

となるから，A が非負定値ならば，すなわち $\alpha_1, \alpha_2, \ldots, \alpha_n$ がすべて非負ならば，任意の $\mathbf{x} \in \mathbf{R}^n$ に対して $A\mathbf{x} \cdot \mathbf{x} \geq 0$ が成り立つ．逆に，$\alpha_1, \alpha_2, \ldots, \alpha_n$ の中に負のものがあれば，$A\mathbf{x} \cdot \mathbf{x} < 0$ となる $\mathbf{x} \in \mathbf{R}^n$ が存在する．例えば $\alpha_j < 0$ ならば $A\mathbf{e}_j \cdot \mathbf{e}_j = \alpha_j < 0$ である．よって，非負定値の場合の主張が得られた．他の場合も同様に考えればよい．

定理 21.3 2変数関数 $f : \mathbf{R}^2 \to \mathbf{R}$ が C^2 級であるとし，点 (a,b) が f の停留点，すなわち，$\dfrac{\partial f}{\partial x}(a,b) = \dfrac{\partial f}{\partial y}(a,b) = 0$ を満たすとする．このとき，(a,b) における f の **Hesse** 行列

$$\nabla^2 f(a,b) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(a,b) & \frac{\partial^2 f}{\partial x \partial y}(a,b) \\ \frac{\partial^2 f}{\partial x \partial y}(a,b) & \frac{\partial^2 f}{\partial y^2}(a,b) \end{pmatrix}$$

が正定値ならば f は (a,b) において極小値を，負定値ならば f は (a,b) において極大値をとり，定値行列でない (正の固有値と負の固有値をもつ) ならば f は (a,b) において極値をとらない．

(証明) f が C^2 級のとき，Taylor の定理より，任意の $k, l \in \mathbf{R}$ に対して

$$f(a+k, b+l) = f(a,b) + \nabla f(a,b) \cdot \begin{pmatrix} k \\ l \end{pmatrix} + \frac{1}{2}\nabla^2 f(a+\theta k, b+\theta l)\begin{pmatrix} k \\ l \end{pmatrix} \cdot \begin{pmatrix} k \\ l \end{pmatrix}$$

と表すことができる．ここで，$\nabla f(a,b) = \begin{pmatrix} \frac{\partial f}{\partial x}(a,b) \\ \frac{\partial f}{\partial y}(a,b) \end{pmatrix}$ であり，θ は $0 < \theta < 1$ を満たすある実数である **(注 1)**．よって，(a,b) が停留点 (すなわち $\nabla f(a,b) = \mathbf{0}$) であって $\nabla^2 f(a,b)$ が正定値ならば，十分小さい k, l に対しては $\nabla^2 f(a+\theta k, b+\theta l)$ も正定値となるから **(注 2)**，前定理より

$$\nabla^2 f(a+\theta k, b+\theta l)\begin{pmatrix} k \\ l \end{pmatrix} \cdot \begin{pmatrix} k \\ l \end{pmatrix} > 0$$

従って

$$f(a+k, b+l) = f(a,b) + \frac{1}{2}\nabla^2 f(a+\theta k, b+\theta l)\begin{pmatrix} k \\ l \end{pmatrix} \cdot \begin{pmatrix} k \\ l \end{pmatrix} > f(a,b)$$

となる．このことは f が (a,b) で極小値をとることを意味する．$\nabla^2 f(a,b)$ が負定値の場合も同様である．また，$\nabla^2 f(a,b)$ が定値行列でないとき，その正の固有値 λ^+ に属する固有ベクトルを $\begin{pmatrix} k^+ \\ l^+ \end{pmatrix}$，負の固有値 λ^- に属する固有ベクトルを $\begin{pmatrix} k^- \\ l^- \end{pmatrix}$ とすれば，やはり上と同様に考えて ($k^\pm, l^\pm$ が十分小さければ) $\quad f(a+k^+, b+l^+) > f(a,b)$ かつ $f(a+k^-, b+l^-) < f(a,b) \quad$ となることがわかる．よって，この場合は f は (a,b) で極値をとらない．

…………………………………………………………………………………………

(注 1) θ は k, l に依存してとられる．

(注 2) $\nabla^2 f(a,b) = \begin{pmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{xy}(a,b) & f_{yy}(a,b) \end{pmatrix}$ が正定値（固有値がともに正）であるための必要十分条件は

「$f_{xx}(a,b) + f_{yy}(a,b) > 0$ かつ $f_{xx}(a,b)f_{yy}(a,b) - f_{xy}(a,b)^2 > 0$」

負定値（固有値がともに負）であるための必要十分条件は

「$f_{xx}(a,b) + f_{yy}(a,b) < 0$ かつ $f_{xx}(a,b)f_{yy}(a,b) - f_{xy}(a,b)^2 > 0$」

定値でない（正の固有値と負の固有値をもつ）ための必要十分条件は

「$f_{xx}(a,b)f_{yy}(a,b) - f_{xy}(a,b)^2 < 0$」

である (確かめよ)．よって，f_{xx}，f_{xy}，f_{yy} の連続性により，a, b の小さな変化に対してこれらの条件が保たれることがわかる．

一般の n 変数関数の場合の Hesse 行列に対しても同様の事実が成り立つが，やや込み入った議論になるのでここでは立ち入らない．