Matrix form for 2^{nd} order Taylor Polynomial

Def: Let $f: \Omega \to \mathbb{R}$ be C^2 ($\Omega \subseteq \mathbb{R}^n$, open).
Then the <u>Hessian matrix</u> of f at $\vec{a} \in \Omega$ is

$$Hf(\vec{a}) = \begin{bmatrix} f_{x_1x_1}(\vec{a}) & \cdots & f_{x_1x_n}(\vec{a}) \\ \vdots & & \vdots \\ f_{x_nx_1}(\vec{a}) & \cdots & f_{x_nx_n}(\vec{a}) \end{bmatrix}$$

<u>Remarks</u> (1) $Hf(\vec{a})$ is $n \times n$ <u>symmetric</u> (by Clairaut's Thm)

(2) In Textbook, Hessian of f = $\det(Hf(\vec{a}))$

So we emphasize our definition is a <u>matrix</u> (More common in advanced level math)

<u>eg</u>: $f(x,y)$ at $(0,0)$

$$Hf(0,0) = \begin{bmatrix} f_{xx}(0,0) & f_{xy}(0,0) \\ f_{yx}(0,0) & f_{yy}(0,0) \end{bmatrix} \qquad (f_{xy} = f_{yx})$$

$$\Rightarrow [x, y] \begin{bmatrix} f_{xx}(0,0) & f_{xy}(0,0) \\ f_{yx}(0,0) & f_{yy}(0,0) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= \underbrace{f_{xx}(0,0)x^2 + 2f_{xy}(0,0)xy + f_{yy}(0,0)y}_{}$$

2^{nd} order term in Taylor polynomials (up to a factor $\frac{1}{2!}$)

2^{nd} order Taylor polynomial of f at $\vec{a}$ in matrix form

$$\boxed{P_2(\vec{x}) = f(\vec{a}) + \vec{\nabla} f(\vec{a})(\vec{x}-\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})^T Hf(\vec{a})(\vec{x}-\vec{a})}$$

where $\vec{\nabla} f(\vec{a})$ regarded as row vector $[f_{x_1}(\vec{a}) \cdots f_{x_n}(\vec{a})]$,

$\vec{x}-\vec{a}$ regarded as column vector $\begin{bmatrix} x_1-a_1 \\ \vdots \\ x_n-a_n \end{bmatrix}$

& $(\vec{x}-\vec{a})^T$ is the transpose $[x_1-a_1 \cdots x_n-a_n]$ (row vector)

eg $g(x,y) = \frac{\ln x}{1-y}$. Find $P_2(x,y)$ at $(1,0)$ using matrix form.

Soln: $g(1,0) = 0$

$$\vec{\nabla} g = \left[\frac{\partial g}{\partial x} \quad \frac{\partial g}{\partial y}\right] = \left[\frac{1}{x(1-y)} \quad \frac{\ln x}{(1-y)^2}\right]$$

$$Hg = \begin{bmatrix} g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{bmatrix} = \begin{bmatrix} -\frac{1}{x^2(1-y)} & \frac{1}{x(1-y)^2} \\ \frac{1}{x(1-y)^2} & \frac{2\ln x}{(1-y)^3} \end{bmatrix}$$

$$\vec{\nabla} g(1,0) = [1 \quad 0]$$

$$Hg = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}$$

$$P_2(x,y) = g(1,0) + \vec{\nabla} g(1,0)\begin{bmatrix} x-1 \\ y \end{bmatrix} + \frac{1}{2}[x-1 \quad y]\, Hg(1,0)\begin{bmatrix} x-1 \\ y \end{bmatrix}$$

$$= 0 + \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x-1 \\ y \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x-1 & y \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x-1 \\ y \end{bmatrix}$$

$$= (x-1) - \frac{1}{2}(x-1)^2 + (x-1)y \qquad \text{(check)}$$

<u>Application to local max/min</u>

If f is C^2, and $\vec{a}$ is a critical point of f.

Then $\vec{\nabla} f(\vec{a}) = \vec{0}$

$$\Rightarrow \quad f(\vec{x}) \approx P_2(\vec{x}) = f(\vec{a}) + \vec{\nabla} f(a)(\vec{x}-\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})^T Hf(a)(\vec{x}-\vec{a})$$

$$= f(\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})^T Hf(a)(\vec{x}-\vec{a})$$

$\therefore$ If $(\vec{x}-\vec{a})^T Hf(a)(\vec{x}-\vec{a}) < 0 \quad \forall \vec{x}$ near $\vec{a}$

then $\quad f(\vec{x}) < f(\vec{a}) \quad \forall \vec{x}$ near $\vec{a}$

$\Rightarrow \quad \vec{a}$ is a <u>local max</u>

If $(\vec{x}-\vec{a})^T Hf(a)(\vec{x}-\vec{a}) > 0 \quad \forall \vec{x}$ near $\vec{a}$

then $\quad f(\vec{x}) > f(\vec{a}) \quad \forall \vec{x}$ near $\vec{a}$

$\Rightarrow \quad \vec{a}$ is a <u>local min</u>

So we need to study when is a sym matrix H satisfies

$$\vec{v}^T H \vec{v} > 0 \quad \forall \text{ vector } \vec{v} \neq \vec{0}$$

and $\quad \vec{v}^T H \vec{v} < 0 \quad \forall$ vector $\vec{v} \neq \vec{0}$

Hence we make the following

Def: Let H be a symmetric $n \times n$ matrix.
Then H is said to be

(1) <u>positive definite</u> if $\vec{x}^T H \vec{x} > 0$
for all column vector $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$

(2) <u>negative definite</u> if $\vec{x}^T H \vec{x} < 0$
for all column vector $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$

(3) <u>indefinite</u> if $\exists$ column vectors $\vec{x}, \vec{y} \in \mathbb{R}^n \setminus \{\vec{0}\}$
such that $\vec{x}^T H \vec{x} > 0$ and $\vec{y}^T H \vec{y} < 0$

<u>Remark</u>: These are <u>not</u> all possibilities : $\exists$ sym. matrix which is not positive definite, negative definite, nor indefinite.

Then the discussion above implies

<u>Thm</u> (<u>Second Derivative Test</u>)

Let $\begin{cases} \bullet\ f: \Omega \to \mathbb{R} \text{ be } C^2,\ \Omega \subseteq \mathbb{R}^n, \text{ open} \\ \bullet\ \vec{a} \in \Omega \text{ such that } \vec{\nabla} f(\vec{a}) = 0 \end{cases}$

Then

$$Hf(\vec{a}) \text{ is } \begin{cases} \underline{\text{positive definite}} \Rightarrow \vec{a} \text{ is a local } \underline{\text{min}} \\ \underline{\text{negative definite}} \Rightarrow \vec{a} \text{ is a local } \underline{\text{max}} \\ \underline{\text{indefinite}} \Rightarrow \vec{a} \text{ is a } \underline{\text{saddle point}} \end{cases}$$

Remark: A critical point which is neither local max nor local min is called a saddle point.

In particular for 2-variable, $\vec{v}^T H \vec{v}$ is of the form

$$g(x,y) = [x \ y]\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + cy^2$$

eg (1) $[x, y]\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = x^2 + 4y^2 > 0 \quad \forall \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \setminus \{\vec{0}\}$

$\therefore \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$ is positive definite

(2) $[x, y]\begin{bmatrix} -1 & 0 \\ 0 & -4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = -x^2 - 4y^2 < 0 \quad \forall \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \setminus \{\vec{0}\}$

$\therefore \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}$ is negative definite

(3) $[x, y]\begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = -x^2 + 4y^2$

If $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $[x, y]\begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = -1 < 0$

If $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $[x, y]\begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 4 > 0$

$\therefore \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}$ is indefinite

(4) $[x,y]\begin{bmatrix}1&0\\0&0\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = x^2 \geq 0 \quad \forall \begin{bmatrix}x\\y\end{bmatrix}\in\mathbb{R}^2\setminus\{\vec{0}\}$

But $[0,1]\begin{bmatrix}1&0\\0&0\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix} = 0 \Rightarrow$ not positive definite

In fact, $\begin{bmatrix}1&0\\0&0\end{bmatrix}$ is not positive definite, negative definite nor indefinite.

(5) $[x,y]\begin{bmatrix}1&2\\2&5\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = x^2+4xy+5y^2$

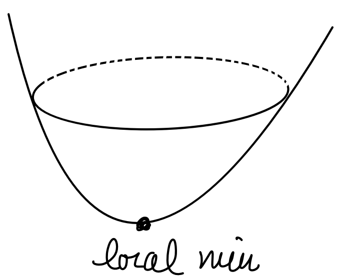
$$\begin{aligned}&= (x^2+4xy+4y^2)+y^2 \\ &= (x+2y)^2+y^2 \\ &> 0 \quad \forall \begin{bmatrix}x\\y\end{bmatrix}\in\mathbb{R}^2\setminus\{\vec{0}\}.\end{aligned}$$

completing square

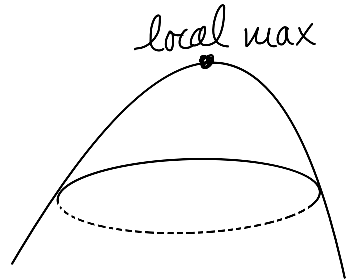
$\therefore \begin{bmatrix}1&2\\2&5\end{bmatrix}$ is positive definite.

<u>Geometrically</u> (locally near the <u>critical point</u>)

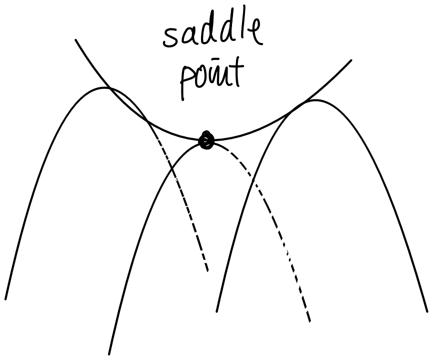
(1) $Hf(\vec{a})$ positive definite



(2) $Hf(\vec{a})$ negative definite



(3) $Hf(\vec{a})$ is indefinite



Thm let $H = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$.

Then

$$H \text{ is } \begin{cases} \underline{\text{positive definite}} \iff \det H = ac - b^2 > 0, \ \& \ a > 0 \\ \underline{\text{negative definite}} \iff \det H = ac - b^2 > 0, \ \& \ a < 0 \\ \underline{\text{indefinite}} \iff \det H = ac - b^2 < 0 \end{cases}$$

Pf: Using completing square

If $a \neq 0$, $[x \ y] \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + cy^2$

$$= a\left(x + \frac{b}{a} y\right)^2 + \frac{(ac - b^2)}{a} y^2$$

(1) $\det H = ac - b^2 > 0 \iff$

$[x \ y] \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ has the same sign as a.

$\therefore$ The 1st 2 statements are proved.

(2) $\det H = ac - b^2 < 0 \iff$ 2 terms have different sign

$\iff \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is indefinite

If $a = 0$, $[x, y] \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2bxy + cy^2 = 2\left(bx + \frac{c}{2}\right) y$

$$= \frac{1}{2}\left[\left(bx + \frac{c}{2}\right) + y\right]^2 - \frac{1}{2}\left[\left(bx + \frac{c}{2}\right) - y\right]^2$$

2 terms of different sign $\Rightarrow$ H indefinite

Also $a = 0 \Rightarrow \det H = ac - b^2 = -b^2 < 0$.

eg1 $q(x,y) = 2xy = [x\ y]\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$,

$(a=0)$ $\det\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = -1$ indefinite

"completing square" $q(x,y) = \frac{1}{2}(x+y)^2 - \frac{1}{2}(x-y)^2$

eg2: $q(x,y) = 17x^2 - 12xy + 8y^2 = [x, y]\begin{bmatrix} 17 & -6 \\ -6 & 8 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$

$17 > 0$ & $\det = 17 \cdot 8 - (-6)^2 = 100 > 0$, positive definite

(completing sq.) $q(x,y) = 17\left(x - \frac{6}{17}y\right)^2 + \frac{100}{17}y^2$

Then for 2-variable, the 2nd derivative test is

Thm (Second Derivative Test for 2-variables)

Let
- $f : \Omega \to \mathbb{R}$ be C^2, $\Omega \subseteq \mathbb{R}^2$, open
- $(a,b) \in \Omega$ such that $\vec{\nabla} f(a,b) = 0$

Then

(1) $f_{xx}f_{yy} - f_{xy}^2 > 0$ & $f_{xx} > 0$ at $(a,b) \Rightarrow (a,b)$ is a local min

(2) $f_{xx}f_{yy} - f_{xy}^2 > 0$ & $f_{xx} < 0$ at $(a,b) \Rightarrow (a,b)$ is a local max

(3) $f_{xx}f_{yy} - f_{xy}^2 < 0$ at $(a,b) \Rightarrow (a,b)$ is a saddle point

(4) $f_{xx}f_{yy} - f_{xy}^2 = 0$ at $(a,b) \Rightarrow$ inconclusive.

Remark $f_{xx}f_{yy} - f_{xy}^2 = \det Hf$ (for 2-variables)

(Some eg s have already been presented in tutorials)

eg 1 $f(x,y) = 3x^2 - 10xy + 3y^2 + 2x + 2y + 3$

Find and classify critical points of f.

Soln: (f polynomial, always C^2)

$$\vec{\nabla} f = \begin{bmatrix} 6x - 10y + 2 & -10x + 6y + 2 \end{bmatrix}$$

$$Hf = \begin{bmatrix} 6 & -10 \\ -10 & 6 \end{bmatrix}$$

Critical point: $\vec{0} = \vec{\nabla} f \Leftrightarrow \begin{cases} 6x - 10y + 2 = 0 \\ -10x + 6y + 2 = 0 \end{cases}$

$$\overset{\text{check}}{\Longleftrightarrow} (x,y) = \left(\tfrac{1}{2}, \tfrac{1}{2}\right)$$

$$f_{xx}f_{yy} - f_{xy}^2 = 6 \cdot 6 - (-10)^2 = -64 < 0$$

$\Rightarrow \left(\frac{1}{2}, \frac{1}{2}\right)$ is a saddle point. (by 2^{nd} derivative test)

(No need to use $f_{xx} = 6 > 0$)

eg 2: $f(x,y) = 3x - x^3 - 3xy^2$

Find and classify critical points of f

Soln (f polynomial, always C^2)

$$\vec{\nabla} f = [3 - 3x^2 - 3y^2 \quad -6xy]$$

$$Hf = \begin{bmatrix} -6x & -6y \\ -6y & -6x \end{bmatrix}$$

Critical point: $\vec{0} = \vec{\nabla} f \iff \begin{cases} 3 - 3x^2 - 3y^2 = 0 \\ -6xy = 0 \end{cases}$

$\overset{\text{check}}{\iff} (x,y) = (0, \pm 1)$ or $(\pm 1, 0)$

(4 critical points)

Critical point	Hf	$f_{xx}f_{yy} - f_{xy}^2$ $= \det Hf$	f_{xx}	classification
$(0,1)$	$\begin{bmatrix} 0 & -6 \\ -6 & 0 \end{bmatrix}$	$-36 < 0$	No need	Saddle point
$(0,-1)$	$\begin{bmatrix} 0 & 6 \\ 6 & 0 \end{bmatrix}$	$-36 < 0$	No need	Saddle point
$(1,0)$	$\begin{bmatrix} -6 & 0 \\ 0 & -6 \end{bmatrix}$	$36 > 0$	$-6 < 0$	local max
$(-1,0)$	$\begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix}$	$36 > 0$	$6 > 0$	local min

eg3 (Inconclusive from 2^{nd} derivative test)

$f(x,y) = x^2+y^4$, $g(x,y)=x^2-y^4$, $h(x,y) = -x^2-y^4$

(at $(0,0)$: local min , saddle , local max)

$\vec{\nabla} f = [2x, 4y^3]$ $\quad \vec{\nabla} g = [2x, -4y^3]$ $\quad \vec{\nabla} h = [-2x, -4y^3]$

($\Rightarrow$ $\vec{\nabla} f(0,0) = \vec{\nabla} g(0,0) = \vec{\nabla} h(0,0) = \vec{0}$)

$Hf = \begin{bmatrix} 2 & 0 \\ 0 & 12y^2 \end{bmatrix}$ $\quad Hg = \begin{bmatrix} 2 & 0 \\ 0 & -12y^2 \end{bmatrix}$ $\quad Hh = \begin{bmatrix} -2 & 0 \\ 0 & -12y^2 \end{bmatrix}$

At the critical point $(0,0)$

$Hf(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} = Hg(0,0)$, $Hh(0,0) = \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix}$

$\Rightarrow$ $\det Hf(0,0) = \det Hg(0,0) = \det Hh(0,0) = 0$

$\therefore$ 2^{nd} derivative test is inconclusive.

Higher dimension example

eg: $q(x,y,z) = xy + yz + zx$

has definite sign for $(x,y,z) \neq (0,0,0)$?

Answer: No

Soln $q = \frac{1}{4}(x+y)^2 - \frac{1}{4}(x-y)^2 + z(x+y)$

Let $u = \frac{x+y}{2}$, $v = \frac{x-y}{2}$, then

$$q = u^2 - v^2 + 2uz$$
$$= (u^2 + 2uz + z^2) - v^2 - z^2$$
$$= (u+z)^2 - v^2 - z^2$$

check

$$= \frac{1}{4}(x+y+2z)^2 - \frac{1}{4}(x-y)^2 - z^2$$

- On the plane $x+y+2z=0$ (i.e. $z = -\frac{x+y}{2}$)

Then $q = q(x, y, -\frac{x+y}{2})$ along the plane

$$= -\frac{1}{4}(x-y)^2 - \frac{1}{4}(x+y)^2$$

< 0 for $(x,y,z) \neq (0,0,0)$ and on the plane

- Along the line $\begin{cases} x-y=0 \\ z=0 \end{cases} \Rightarrow \begin{matrix} x=y \\ z=0 \end{matrix}$

$q = q(x,x,0) = \frac{1}{4}(x+x+0)^2 - 0 - 0 = x^2 > 0$ for $x \neq 0$

$(x,x,0)$ on the line

Together $\Rightarrow (0,0)$ is a saddle point.

Second Derivative Test for general n

Recall f is $C^2 \Rightarrow$ (by Clairaut's / mixed derivative Thm)

$$Hf(\vec{a}) = \left[f_{x_i x_j} \right]_{i,j=1,\cdots,n} \text{ is symmetric}$$

Theory of Linear Algebra $\Rightarrow$ Hf is diagonizable

i.e. $\exists$ orthogonal $n \times n$ matrix P (i.e. $P^T P = Id$) s.t.

$$P^T Hf(\vec{a}) P = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

where λ_i, $i = 1, \cdots, n$, are eigenvalues of $Hf(\vec{a})$.

$$\Rightarrow \boxed{ Hf(\vec{a}) \text{ is } \begin{cases} \text{positive definite} \Leftrightarrow \text{all } \lambda_i > 0 \\ \text{negative definite} \Leftrightarrow \text{all } \lambda_i < 0 \\ \text{indefinite} \Leftrightarrow \text{some } \lambda_i > 0, \text{ some } \lambda_j < 0 \ (\text{all} \neq 0) \end{cases} }$$

Another way to check is consider determinants of submatrix

For each $1 \leq k \leq n$,

consider submatrix H_k given by the upper left $k \times k$ entries.

$$\begin{bmatrix} f_{x_1 x_1} & \cdots & f_{x_1 x_k} & \cdots & f_{x_1 x_n} \\ \vdots & & \vdots & & \vdots \\ f_{x_k x_1} & \cdots & f_{x_k x_k} & \cdots & f_{x_k x_n} \\ \vdots & & \vdots & & \vdots \\ f_{x_n x_1} & \cdots & f_{x_n x_k} & \cdots & f_{x_n x_n} \end{bmatrix}$$

Then

$$\boxed{\begin{array}{l} Hf(\vec{a}) \text{ is positive definite} \iff \det H_k > 0, \ \forall k = 1, \cdots, n \\ Hf(\vec{a}) \text{ is negative definite} \iff \det H_k \begin{cases} < 0, & k \text{ odd} \\ > 0, & k \text{ even} \end{cases} \end{array}}$$

<u>egs</u> (1) $n=2$ $\qquad \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$ has $\quad H_1 = [f_{xx}]$

$$H_2 = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

$$\Rightarrow \quad \det H_1 = f_{xx}$$

$$\det H_2 = f_{xx} f_{yy} - f_{xy}^2$$

(Same result as before)

(2) Diagonal matrix $\begin{bmatrix} \lambda_1 & & & & 0 \\ & \ddots & & & \\ & & \lambda_k & & \\ & & & \ddots & \\ 0 & & & & \lambda_n \end{bmatrix} \Rightarrow H_k = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k \end{bmatrix}$

$$\Rightarrow \det H_k = \lambda_1 \cdots \lambda_k .$$