EIGENVALUES AND EIGENVECTORS

$Ax = \lambda X$ λ = eigenvalue, X = eigenvector

$(A-\lambda I)X = 0$ with $X \neq 0$ $\Rightarrow$ A singular $\Leftrightarrow$ $\det(A-\lambda I) = 0$

$\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ $\sigma(A)$ = spectrum (λ, X) = eigenpair

$X \in N(A-\lambda I)$ $N(A-\lambda I)$ = eigenspace

For 2×2 matrices: $\sum_{i=1}^{n} \lambda_i = \text{Trace}(A)$ $\prod_{i=1}^{n} \lambda_i = \det(A)$

The eigenvalues of A are the same of A^T

$\frac{1}{\lambda} = \lambda^{-1}$ is an eigenvalue of A^{-1} $\rho(A) := \max\{|\lambda_i|\}$ $\rho(A)$ = spectral radius

SIMILARITY

<u>Goal</u>: transform into another matrix of easier structure (tri. or diag.)

$S^{-1}AS = B$ $S^{-1}AS$ = similarity transformation

If A $n \times n$ with n linearly independent eigenvectors $\{x_1, \dots, x_n\}$

$S^{-1}AS = \Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ n distinct eigenvalues = n dist. e.vectors

ALGEBRAIC AND GEOMETRIC MULTIPLICITY

Alg.mult. = number of times an eigenvalue is a root of the ch. polynomial.

Geo.mult. = dimension of the eigenspace $\dim(N(A-\lambda I))$

<u>In general</u>, geo.mult $\leq$ alg.mult geo.mult = alg.mult $\rightarrow$ λ_i semi-simple

When eigenvalues are all distinct $\Rightarrow$ λ_i simple

simple or semi-simple $\Rightarrow$ set of linearly independent eigenvectors

DIAGONALIZABILITY

Question: when A is similar to a diag. matrix? $\Rightarrow$ A: complete set of e.vect.

JORDAN CANONICAL FORM

$J = \text{diag}\left(J(\lambda_1), J(\lambda_2), \dots, J(\lambda_n)\right)$ $\quad J(\lambda_i)$ = Jordan segments

$J(\lambda_i) := \begin{pmatrix} J_1(\lambda_i) & & 0 \\ & \ddots & \\ 0 & & J_{t_i}(\lambda_i) \end{pmatrix}$ $\quad t_i :=$ geo.mult. of λ_i

$\sigma(A)$ = diag. elements of J

$J_i(\lambda_j) = \begin{pmatrix} \lambda_j & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_j \end{pmatrix}$ $\quad \#\lambda_j$ = alg.mult of λ_j

Not similar to diagonal matrix $\Rightarrow$ defective matrix $\Rightarrow$ nilpotent matrix

Given $N^{n\times n}$ with $N^k = 0$ with $k>0$ $\quad N$ = nilpotent

CORE-NILPOTENT DECOMPOSITION

A singular matrix of index k such that $\text{rank}(A^k) = r$

$Q^{-1} A Q = \begin{pmatrix} C & 0 \\ 0 & N \end{pmatrix}$ $\quad C$ = non-singular, N = nilpotent of index k

Index = smallest non negative integer k such that $\text{rank}(A^k) = \text{rank}(A^{k+1})$

How to construct a core-nilpotent decomposition for 2×2

1. Determine $\text{index}(A) = k$
2. Compute $\text{rank}(A^k) = r$
3. Construct $Q = (X|Y)$ with $X = R(A^k)$, $Y = N(A^k)$
4. Compute Q^{-1}
5. Compute $Q^{-1} A Q$ to verify

$A^T A = A A^T$ = normal m.

ORTHOGONAL MATRICES

$Q \in \mathbb{R}^{n \times n}$ orthogonal $\Rightarrow Q^TQ = QQ^T = I \Rightarrow Q^{-1} = Q^T \Rightarrow$ isometries
$\Rightarrow$ unitary m. if $\mathbb{C}^{n \times n}$

$\|x\|_2 = \sqrt{x^Tx} = \sqrt{\sum x_i^2} \qquad \|Qx\|_2 = \|x\|_2$

ELEMENTARY ORTHOGONAL PROJECTORS

$u \in \mathbb{R}^n \quad \|u\|_2 = 1 \quad Q = I - uu^T \quad Q$ = elementary projector (orthogonal)

In general, if $\|u\|_2 \neq 1 \Rightarrow Q = I - \frac{uu^T}{u^Tu} = I - \frac{uu^T}{\|u\|_2^2}$

ELEMENTARY REFLECTORS (HOUSEHOLDER T.)

$u \in \mathbb{R}^n \quad R = I - 2\frac{uu^T}{u^Tu} \quad R$ = elementary reflector

If $\|u\|_2 = 1 \Rightarrow R = I - 2uu^T$

ORTHOGONAL REDUCTION (THEOREM)

For every matrix $A \in \mathbb{C}^{m \times n}$ ($\mathbb{R}^{m \times n}$) there exists a unitary matrix U and an orthogonal and upper trapezoidal matrix T such that $A = UT$

$\Rightarrow$ QR factorization

GIVENS ROTATION

$G = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \quad c^2 + s^2 = 1 \quad G$ orthogonal

$z = \begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow c = \frac{x}{\sqrt{x^2+y^2}} = \frac{x}{\|z\|_2}, \quad s = \frac{y}{\sqrt{x^2+y^2}} = \frac{y}{\|z\|_2} \Rightarrow W = \begin{pmatrix} \|z\|_2 \\ 0 \end{pmatrix}$

LEAST SQUARE PROBLEM

$\min_x \|Ax-b\|_2^2$ $\quad A = QR$ with $Q=(Q_1\, Q_2)$ $\quad Rx = Q_1 b \Rightarrow Q^TA = R$

$$Q^T(A|B) = (Q^TA \mid Q^TB) = \left(\begin{pmatrix} R \\ 0 \end{pmatrix} \middle| \begin{pmatrix} Q_1^T b \\ Q_2^T b \end{pmatrix}\right) \quad \text{Solve } Rx = Q_1^T b$$

COMPLEMENTARY SUBSPACES

$\mathcal{X}, \mathcal{Y}$ subspaces $\quad \mathcal{X}, \mathcal{Y} \subseteq V$ $\qquad \mathcal{X}, \mathcal{Y}$ complementary $\Rightarrow V = \mathcal{X} \oplus \mathcal{Y}$

$\oplus$ = direct sum $\qquad \mathcal{X} \cap \mathcal{Y} = \emptyset$

$\mathcal{X} = \text{span}\{x_1, \dots, x_r\}$ $\quad \mathcal{Y} = \text{span}\{y_1, \dots, y_p\}$

$V = \text{span}\{x_1, \dots, x_r, y_1, \dots, y_p\}$

x = projection of v onto $\mathcal{X}$ along $\mathcal{Y}$

y = projection of v onto $\mathcal{Y}$ along $\mathcal{X}$

$v = x + y \quad x \in \mathcal{X} \quad y \in \mathcal{Y} \quad v \in V$

ORTHOGONAL COMPLEMENTARY SUBSPACE

$\langle x | y \rangle = 0 \Rightarrow x, y$ orthogonal $\Rightarrow x^Ty = 0$

Let $M \subseteq V$ a subset of $V \Rightarrow x \in M$, $y \in M^\perp$, $x^Ty = 0 \Rightarrow M \oplus M^\perp = V$

$R(A) \perp N(A^T)$ $\qquad V = \mathbb{R}^m = R(A) \oplus N(A^T)$

$R(A^T) \perp N(A)$ $\qquad V = \mathbb{R}^n = R(A^T) \oplus N(A)$

QR factorization $\Rightarrow$ information about $R(A), N(A^T)$

URV factorization $\Rightarrow$ information about $R(A), R(A^T), N(A), N(A^T)$

URV FACTORIZATION

Given $A \in \mathbb{R}^{m \times n}$ with $\text{rank}(A) = r$ there exist two orthogonal matrices $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ and a non-singular matrix $C \in \mathbb{R}^{r \times r}$ such that

$$A = U \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} V^T = URV^T$$

1. The first r columns of U are an orthonormal basis for $R(A)$
2. The first r columns of V are an orthonormal basis for $R(A^T)$
3. The last m-r columns of U are an orthonormal basis for $N(A^T)$
4. The last n-r columns of V are an orthonormal basis for $N(A)$

$$U = (\underbrace{u_1, \dots, u_r}_{\text{b. } R(A)}, \underbrace{\dots, u_m}_{\text{b. } N(A^T)}) \qquad V = (\underbrace{v_1, \dots, v_r}_{\text{b. } R(A^T)}, \underbrace{\dots, v_n}_{\text{b. } N(A)})$$

SINGULAR VALUE DECOMPOSITION

Any matrix $A \in \mathbb{R}^{m \times n}$ with $\text{rank}(A) = r$ can be factorized as

$$A = U \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} V^T = U \Sigma V^T \qquad D = \text{diag}(\sigma_1, \dots, \sigma_r)$$

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq 0$$

U = left-hand singular vectors

V = right-hand singular vectors

$\# \sigma_i = \text{rank}(A) \Rightarrow UU^T = U^TU = I$

$VV^T = V^TV = I$

$$A = (U_A \; U_A^{\perp}) \Sigma \begin{pmatrix} V_A^T \\ (V_A^{\perp})^T \end{pmatrix}$$

$\|A\|_2 = \sigma_1$ = 1st singular value = $\sqrt{\lambda_{max}}$

ECKART-YOUNG (THEOREM)

Let $A \in \mathbb{R}^{m \times n}$ with $\text{rank}(A) = r$ and fix $k < r$, then the matrix approx. problem $\min \|A - Z\|_2$ with $\text{rank}(Z) = k$ has the solution

$$Z = A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T$$

and the value of the minimum is $\|A - A_k\|_2 = \sigma_{k+1}$

RELATION WITH EIGENVALUES

$A = U \Sigma V^T$

$$A^T A = V \Sigma^T \underbrace{U^T U}_{I} \Sigma V^T = V \Sigma^2 V^T = V \Lambda V^T$$

$$A A^T = U \Sigma \underbrace{V^T V}_{I} \Sigma^T U^T = U \Sigma \Sigma^T U^T = U \Sigma^2 U^T = U \Lambda U^T$$

$$\sigma(A^T A) = \{\sigma_1, \dots, \sigma_r\} = \sigma(A A^T)$$

PRINCIPAL COMPONENT ANALYSIS

Goal: q principal axis that return maximal variance under projection

We assume to have centred data $x_j = t_j - \text{mean}(t_j)$ $j = 1, \dots, N$

$\Rightarrow \text{mean}(x_j) = 0 \Rightarrow \text{var}(x_j) = \dfrac{x_j^T x_j}{m} \Rightarrow \text{std}(x_j) = \dfrac{x_j}{\sqrt{\text{var}(x_j)}}$

$X = (x_1 | x_2 | \dots | x_N)$ $\quad X \in \mathbb{R}^{m \times n}$ $\quad \Rightarrow$ (SVD) $X = U \Sigma V^T$ with $\text{rank}(X) = r$

v_i = principal component directions

u_i = normalized principal components

$z_1 = X v_1 \Rightarrow$ vector with largest sample variance among all lin. comb. of X

$z_1 \in R(X)$

Covariance matrix $\Rightarrow \dfrac{X^T X}{m-1}$ $\qquad x_j = \dfrac{t_j - \mu_{t_j}}{\text{std}(t_j)} \Rightarrow X^T X$ correlation matrix

PROJECTION MATRIX

$X, Y \subseteq V$ subspaces $\quad V = X \oplus Y \quad v \in V \Rightarrow v = x + y \quad x \in X, y \in Y$

$Pv = x \quad P =$ projection matrix $\Rightarrow Pv = x$ projection of v onto X along Y

$X = \text{span}\{x_1, \dots, x_r\} \quad Y = \text{span}\{y_1, \dots, y_{m-r}\}$

$B = \{x_1, \dots, x_r, y_1, \dots, y_{m-r}\} \quad \text{rank}(B) = m \Rightarrow B = (X|Y)$

$PB = P(X|Y) = (PX|PY) = (X|0) \Rightarrow PBB^{-1} = (X|0)B^{-1} \Rightarrow P = (X|0)B^{-1}$

$\Rightarrow P = B\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}B^{-1}$ projector onto X along Y

$I - P = B\begin{pmatrix} 0 & 0 \\ 0 & I_r \end{pmatrix}B^{-1}$ projector onto Y along X

$R(P) = \{x \mid Py = x\} = X \quad N(P) = R(I-P) \quad \|P\|_2 = 1$

From SVD or URV factorization:

$P_{R(A)} = UU^T \quad P_{R(A)^\perp} = P_{N(A^T)} = I - P_{R(A)}$

$P_{R(A^T)} = VV^T \quad P_{R(A^T)^\perp} = P_{N(A)} = I - P_{R(A^T)}$

$P_{R(A)} = M(M^TM)^{-1}M^T$

$M = R(A)$

MULTIDIMENSIONAL CALCULUS

$f(x,y)$ = function of two variables = surface/plane

$A^TAx = A^Tb \Rightarrow$ normal equation $\Rightarrow$ obtained by partial derivatives

$f(x,y) \Rightarrow \frac{\partial f}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x} = f'_x \qquad f(x) \Rightarrow f'(x) \Rightarrow \frac{df}{dx}$

$\frac{\partial f}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y+\Delta y) - f(x,y)}{\Delta y} = f'_y$

second partial derivatives $\Rightarrow \frac{\partial}{\partial x}\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial x^2} \quad \frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial y \partial x}$

HESSIAN MATRIX

$$H = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

$H = H^T \Rightarrow$ symmetric

The Hessian matrix of $f(x) = \|Ax - b\|_2^2$ is equal to

$$H = 2A^TA$$

If $f(x)$ with $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ then $H_{ij} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{ij}$

The vector that collect all the first derivatives (partial) is the

GRADIENT VECTOR ∇f

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

TANGENT LINE

$y = f(x_0) + f'(x_0)(x - x_0)$ $\quad x_0$ = stationary point if $f'(x_0) = 0$

Taylor polynomial $\Rightarrow f(x) = f(x_0) + \cancel{f'(x_0)(x-x_0)} + \frac{f''(x_0)}{2}(x-x_0)^2 + \dots$

$$f(x) = f(x_0) + \frac{f''(x_0)}{2}(x-x_0)^2 + \dots$$

If $f''(x_0) > 0$ then $f(x) \geq f(x_0) \Rightarrow x_0$ minimum point

TANGENT PLANE

$f(x,y) = x^2y^2 + xy + y$ $\quad \frac{\partial f}{\partial x} = 2xy^2 + y$ $\quad \frac{\partial f}{\partial y} = 2x^2y + x + 1$

$$z = x_0^2 y_0^2 + x_0 y_0 + y_0 + (2x_0 y_0^2 + y_0)(x - x_0) + (2x_0^2 y_0 + x_0 + 1)(y - y_0)$$

z is the tangent plane to (x_0, y_0).

NORMAL VECTOR

$\|z\|=1$ vector of 3 components perpendicular to the tangent plane

$$\left(\frac{\partial f}{\partial x}\right)_0 = \frac{\partial f}{\partial x}(x_0,y_0) \qquad \left(\frac{\partial f}{\partial y}\right)_0 = \frac{\partial f}{\partial y}(x_0,y_0)$$

$$N = \begin{pmatrix} \left(\frac{\partial f}{\partial x}\right)_0 \\ \left(\frac{\partial f}{\partial y}\right)_0 \\ -1 \end{pmatrix}$$

$$z = z_0 + \left(\frac{\partial f}{\partial x}\right)_0 (x-x_0) + \left(\frac{\partial f}{\partial y}\right)_0 (y-y_0)$$

$z_0 = f(x_0,y_0)$

z = tangent plane

$$\begin{pmatrix} \left(\frac{\partial f}{\partial x}\right)_0 (x-x_0) \\ \left(\frac{\partial f}{\partial y}\right)_0 (y-y_0) \\ -(z-z_0) \end{pmatrix} = \left(\left(\frac{\partial f}{\partial x}\right)_0, \left(\frac{\partial f}{\partial y}\right)_0, -1\right) \begin{pmatrix} (x-x_0) \\ (y-y_0) \\ \vdots \\ (z-z_0) \end{pmatrix} = 0$$

if (x_0,y_0) is a point on the tangent plane

We need to normalize N by $\left\| \left(\frac{\partial f}{\partial x}\right)_0, \left(\frac{\partial f}{\partial y}\right)_0, -1 \right\|_2$

DIFFERENTIALS

$dx = (x-x_0)$

$dy = (y-y_0)$

$dz = (z-z_0)$

$$dz = \left(\frac{\partial f}{\partial x}\right)_0 dx + \left(\frac{\partial f}{\partial y}\right)_0 dy \quad \Rightarrow \quad z_0 = f(x_0,y_0)$$

$$(df)_0 = \left(\frac{\partial f}{\partial x}\right)_0 dx + \left(\frac{\partial f}{\partial y}\right)_0 dy \qquad (x_0,y_0) \text{ fixed}$$

$$(df)_0 = z - z_0 = f(x,y) - f(x_0,y_0)$$

$df = f_x dx + f_y dy$ is called total differential

We are approximating $\Delta f \approx f_x \Delta x + f_y \Delta y$

To find the minimum of a function $f(x)$ we need to find the zeros of ∇f

If $g = \nabla f$, the first partial derivatives of g are the second partial derivatives of f

The Jacobian matrix J_g collect all the first derivatives (partial)

$$(J_g)_{ij} = \left(\frac{\partial g_i}{\partial x_j}\right) \qquad (J_g) = (H_f) \text{ if } g = \nabla f$$

NEURAL NETWORK

$\sigma(x) = \frac{1}{1+e^{-x}}$ $\quad\sigma'(x) = \sigma(x)(1-\sigma(x))$

$\tan\{h\} = \frac{1-e^{2x}}{1+e^{2x}}$ $\quad \tanh' = \frac{2}{(1+e^{-x})^2}$

These are activation functions for Neural Networks

We can modify those functions applying shifting and stretching

$\sigma(3(x-5))$ $\quad 3 \Rightarrow$ scaling $\Rightarrow$ weighting $\Rightarrow w_i$

$5 \Rightarrow$ shifting $\Rightarrow$ biasing $\Rightarrow b_i$

$z_i = W_i x + b_i \qquad z_i = W_{ij} x_j + b_i \Rightarrow z = Wa + b \Rightarrow \sigma(z) = \sigma(Wa+b)$

If we are using more values, the number of nested functions will correspond to the number of layers of the network.

If we use one value, we have the perceptron (simplest neural network)

$F(x)$ = steps of the neural network

$$y(x^{[i]}) = \begin{cases} \binom{1}{0} & \text{if } x^{[i]} \text{ in category A} \\ \binom{0}{1} & \text{if } x^{[i]} \text{ in category B} \end{cases}$$

$y(x^{[i]})$ multivariate function of the neural network

$cost(W^{[i]}, b^{[i]})$ = error cost function $\qquad cost = \frac{1}{N}\sum_{i=1}^{N} \| y(x^{\{i\}}) - F(x^{\{i\}}) \|_2^2$

To compute $F(x) \Rightarrow \min\limits_{W^{\{i\}}, b^{\{i\}}} cost$

DIRECTIONAL DERIVATIVE $D_u f$ $u=(u_1,u_2)$ direction

$$D_u f = \frac{\partial f}{\partial x}u_1 + \frac{\partial f}{\partial y}u_2$$

GRADIENT $\nabla f = \frac{\partial f}{\partial x}i + \frac{\partial f}{\partial y}j$ (add $\frac{\partial f}{\partial z}k$ in 3D)

The gradient tells the direction to climb and its length $|\nabla f|$ gives the steepness

$f = 3x+y+1$ $\nabla f = 3i+j$ $\nabla f = \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix}$

CHAIN RULE $f(g(x,y))$ $\frac{\partial f}{\partial x} = \frac{df}{dg}\frac{\partial g}{\partial x}$ $\frac{\partial f}{\partial y} = \frac{df}{dg}\frac{\partial g}{\partial y}$

1. The leadings candidates are stationary points ($\frac{\partial f}{\partial x}=0$)
2. The other candidates are rough points (no derivative) and endpoints
3. Maximum or minimum is decided by the sign of second derivative

If derivative exists at an interior min or max, it's zero:

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = 0 \qquad \frac{\partial f}{\partial x}=0, \quad \frac{\partial f}{\partial y}=0$$

At (0,0) the quadratic function $f(x,y) = ax^2+bxy+cy^2$ has a

min if $a>0$, $ac>b^2$ maximum if $a<0$, $ac>b^2$ saddle if $ac<b^2$

$\boxed{\text{CURVE}}$ $y=f(x)$ $\boxed{\text{SURFACE}}$ $z=f(x,y)$

$\frac{df}{dx} \Rightarrow \frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ $\quad \frac{d^2f}{dx^2} \Rightarrow \frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y^2}, \frac{\partial^2 f}{\partial y \partial x}$

TANGENT LINE $\Rightarrow$ $\boxed{\text{TANGENT PLANE}}$ $z - z_0 = \frac{\partial f}{\partial x}(x-x_0) + \frac{\partial f}{\partial y}(y-y_0)$

$$\frac{\partial f}{\partial x}(x,y) = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x}$$

$f(x,y) = x^2y^2 + xy + y \Rightarrow \frac{\partial f}{\partial x}(x,y) = 2xy^2 + y$ $\boxed{\text{PARTIAL DERIVATIVES}}$

$f(x,y) = y^2 - x^2 \Rightarrow \frac{\partial f}{\partial x} = \cancel{y^2} - 2x \quad \frac{\partial f}{\partial y} = 2y$

$\boxed{\text{SADDLE POINT}}$ $\Rightarrow$ The bottom of the y parabola at (0,0) is the top of the x parabola $\Rightarrow$ neither max nor min!

$z = 14 - x^2 - y^2$ find tangent plane at $(x_0, y_0, z_0) = (1, 2, 9)$

$\left(\frac{\partial f}{\partial x}\right)_0 = -2x_0 = -2(1) = -2$

$\left(\frac{\partial f}{\partial y}\right)_0 = -2y_0 = -2(2) = -4$

$z - z_0 = \left(\frac{\partial f}{\partial x}\right)_0 (x-x_0) + \left(\frac{\partial f}{\partial y}\right)_0 (y-y_0)$

$z - 9 = (-2)(x-1) + (-4)(y-2) =$

$= -2x + 2 - 4y + 8 =$

$= -2x - 4y + 10$

$z = -2x - 4y + 19$

$N = \begin{pmatrix} \left(\frac{\partial f}{\partial x}\right)_0 \\ \left(\frac{\partial f}{\partial y}\right)_0 \\ -1 \end{pmatrix}$ NORMAL VECTOR

$\boxed{\text{NEWTON'S METHOD}}$ $\Rightarrow$

$\left(\frac{\partial g}{\partial x}\right)\Delta x + \left(\frac{\partial g}{\partial y}\right)\Delta y = -g(x_n, y_n)$ $\quad \Delta x, \Delta y \rightarrow$

$\left(\frac{\partial h}{\partial x}\right)\Delta x + \left(\frac{\partial h}{\partial y}\right)\Delta y = -h(x_n, y_n)$ $\quad (x_n, y_n) \rightarrow (x_{n+1}, y_{n+1})$