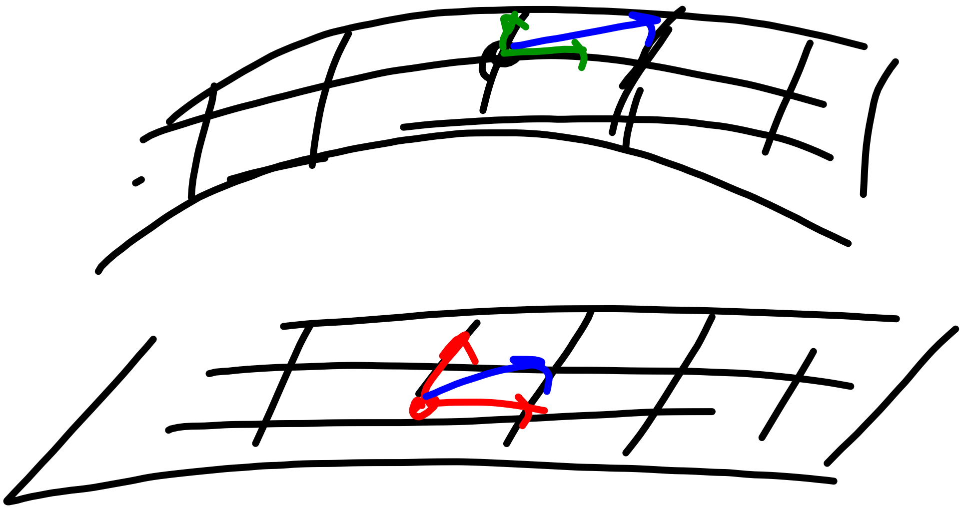

$\mathbb{R}^n$.

PARTIAL DERIVS DETERMINE RATE OF CHANGE AS YOU MOVE IN THE DIRECTION OF GRID LINES IN THE DOMAIN

- WE NOW DEFINE A DERIVATIVE WHEN MOVING IN ANY DIRECTION
- THE GRADIENT TELLS US THE DIRECTION OF GREATEST INCREASE

THEOREM: IF f IS DIFFERTIABLE AT $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ AND $\underline{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$

THEN

$$\frac{\partial f}{\partial \underline{v}}(\underline{x}) = \nabla f(\underline{x}) \cdot \underline{v}.$$

IN PARTICULAR

$$\frac{\partial f}{\partial (2 \cdot \underline{v})}(\underline{x}) = 2 \cdot \frac{\partial f}{\partial \underline{v}}(\underline{x}).$$

(REPLACE $\nabla f(\underline{x}) \cdot \underline{v}$ WITH $\nabla f(\underline{x}) \cdot (2\underline{v})$.)

SINCE f IS DIFFERTIABLE AT $\underline{x}$,

$$\lim_{\underline{y}\to\underline{x}} \frac{f(\underline{y})-f(\underline{x})-\nabla f(\underline{x})\cdot(\underline{y}-\underline{x})}{\|\underline{y}-\underline{x}\|}=0.$$

CHOOSE $\underline{y}=\underline{x}+t\underline{v}$, LET $t\to 0$.

$$\lim_{t\to 0} \frac{f(\underline{x}+t\underline{v})-f(\underline{x})-\nabla f(\underline{x})\cdot t\underline{v}}{\|t\underline{v}\|}=0.$$

$\|t\underline{v}\|=|t|\cdot\|\underline{v}\|.$

THUS

$$\lim_{t\to 0}\frac{f(\underline{x}+t\underline{v})-f(\underline{x})}{t}=\lim_{t\to 0}\nabla f(\underline{x})\cdot\underline{v}$$

$$=\nabla f(\underline{x})\cdot\underline{v}. \qquad \square$$

THEOREM: IF $\nabla f(\underline{x}) \neq \underline{0}$,

$\frac{\nabla f}{\|\nabla f\|} = \underline{v}$ MAXIMIZES

$\frac{\partial f}{\partial \underline{v}}(\underline{x})$ AMONG ALL VECTORS

$\underline{v}$ WITH $\|\underline{v}\| = 1$.

"THE GRADIENT IS THE DIRECTION OF MAXIMUM INCREASE."

<u>EXAMPLE:</u> MAXIMIZE SOME UNKNOWN FUNCTION f, WHOSE VALUE CAN BE SAMPLED FROM A BLACK BOX.

START WITH SOME INITIAL GUESS $\underline{x}_0$. CALCULATE

$$\nabla f(x_0) \approx \frac{1}{\delta} \begin{pmatrix} f(\underline{x}_0 + \delta e_1) - f(\underline{x}_0) \\ f(\underline{x}_0 + \delta e_2) - f(\underline{x}_0) \\ \vdots \\ f(\underline{x}_0 + \delta e_n) - f(\underline{x}_0) \end{pmatrix}$$

UPDATE TO

$$\underline{x}_1 = \underline{x}_0 + c \cdot \frac{\nabla f(\underline{x}_0)}{\|\nabla f(\underline{x}_0)\|}.$$

ITERATE.

THIS METHOD IS <u>GRADIENT ASCENT</u>.

IF $\mathbb{R}^2 \xrightarrow{f} \mathbb{R}$, GRAPHED IN $\mathbb{R}^3$ AS

$$\text{GRAPH}_f(\underline{x}) = \begin{pmatrix} \underline{x} \\ f(\underline{x}) \end{pmatrix}.$$

$\underline{u} \in \mathbb{R}^2$ A UNIT VECTOR,

$\frac{\partial f}{\partial u}(\underline{x})$ IS THE SLOPE OF TANGENT LINE AT $(\underline{x}, f(\underline{x}))$ AS $\underline{x}$ MOVES IN DIRECTION $\underline{u}$.

<u>THEOREM:</u> LET $f: \mathbb{R}^n \to \mathbb{R}$.
BE DIFFERENTIABLE ON AN OPEN SET CONTAINING THE SEGMENT FROM $\underline{x}$ TO $\underline{y}$. THEN THERE IS AN $\underline{x}_0$ ON THIS SEGMENT SUCH THAT

$$f(\underline{y}) - f(\underline{x}) = \nabla f(\underline{x}_0) \cdot (\underline{y} - \underline{x}).$$

<u>DEFINITION:</u> WE SAY AN OPEN SET U OF $\mathbb{R}^n$ IS <u>POLYGONALLY CONNECTED</u> IF, GIVEN ANY TWO POINTS $\underline{x}, \underline{y} \in U$, THERE IS A SEQUENCE OF POINTS $\underline{x} = \underline{x}_0, \underline{x}_1, \underline{x}_2, \dots, \underline{x}_n = \underline{y}$ SO THAT, FOR EACH i, THE LINE SEGMENT FROM $\underline{x}_i$ TO $\underline{x}_{i+1}$ IS CONTAINED IN U.

THEOREM: IF $f : \mathbb{R}^n \to \mathbb{R}$ IS DIFF ON A POLYGONALLY CONNECTED, OPEN SET AND $\nabla f = 0$, THEN f IS CONSTANT ON THE SET.

DEFINITION: A VECTOR VALUED FUNCTION $f: \mathbb{R}^n \to \mathbb{R}^m$ IS **DIFFERENTIABLE** IF EACH COORDINATE FUNCTION

$$f_1, \ldots, f_m : \mathbb{R}^n \to \mathbb{R}$$

IS DIFFERENTIABLE.

IT IS CONVENIENT IN THIS SITUATION TO WRITE THE GRADIENT OF THE COMPONENT FUNCTIONS AS ROW VECTORS

$$\nabla f_1(\underline{x}) = \left(\frac{\partial f_1}{\partial x_1}, \frac{\partial f_1}{\partial x_2}, \dots, \frac{\partial f_1}{\partial x_n}\right)$$

$$\nabla f_2(\underline{x}) = \left(\frac{\partial f_2}{\partial x_1}, \frac{\partial f_2}{\partial x_2}, \dots, \frac{\partial f_2}{\partial x_n}\right)$$

$$\vdots$$

$$\nabla f_m(\underline{x}) = \left(\frac{\partial f_m}{\partial x_1}, \frac{\partial f_m}{\partial x_2}, \dots, \frac{\partial f_m}{\partial x_n}\right).$$

IN THIS WAY, THE DOT PRODUCT

$$D_{\underline{v}} f(\underline{x}) = \nabla f(\underline{x}) \cdot \underline{v}$$

IS <u>MATRIX</u> <u>MULTIPLICATION</u>.

THE DERIVATIVE MATRIX

IS THE MATRIX WITH GRADIENTS OF COMPONENT FUNCTIONS AS ROW VECTORS.

$$Df(\underline{x}) = f'(\underline{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & & & \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

EXAMPLE:

$$f(x,y) = \begin{pmatrix} x^2+2xy+y^2 \\ xy^2+yx^2 \end{pmatrix}$$

$$f'(x,y) = \begin{pmatrix} 2x+2y & 2x+2y \\ y^2+2xy & 2xy+x^2 \end{pmatrix}$$

DEFINITION: THE DEGREE 1 TAYLOR POLYNOMIAL OF $f: \mathbb{R}^n \to \mathbb{R}^m$ AT $\underline{x}_0$ IS

$$T(\underline{x}) = f(\underline{x}_0) + f'(\underline{x}_0)(\underline{x} - \underline{x}_0).$$

EXAMPLE: $f(x,y,z) = \begin{pmatrix} -x+y+z \\ 2x-2y+2z \\ 3x+3y-3z \end{pmatrix}$

$$f'(x,y,z) = \begin{pmatrix} -1 & 1 & 1 \\ 2 & -2 & 2 \\ 3 & 3 & -3 \end{pmatrix}.$$

PROOF: f DIFF AT $\underline{x}_0$

$\iff f_1, \dots, f_m$ DIFF AT $\underline{x}_0$

$$\iff \lim_{\underline{x} \to \underline{x}_0} \frac{f_i(\underline{x}) - f_i(\underline{x}_0) - \nabla f_i(\underline{x}_0)(\underline{x} - \underline{x}_0)}{\|\underline{x} - \underline{x}_0\|} = 0,$$

$$\iff \lim_{\underline{x} \to \underline{x}_0} \frac{f(\underline{x}) - f(\underline{x}_0) - f'(\underline{x}_0)(\underline{x} - \underline{x}_0)}{\|\underline{x} - \underline{x}_0\|} = \underline{0},$$

SINCE EACH COORD CONVERGES

IF A IS ANY OTHER MATRIX WRITE $A = f'(\underline{x}_0) + B$.

$$\lim_{\underline{x} \to \underline{x}_0} \frac{f(\underline{x}) - f(\underline{x}_0) - A(\underline{x} - \underline{x}_0)}{\|\underline{x} - \underline{x}_0\|}$$

$$= \lim_{\underline{x} \to \underline{x}_0} \frac{f(\underline{x}) - f(\underline{x}_0) - f'(\underline{x}_0)(\underline{x} - \underline{x}_0)}{\|\underline{x} - \underline{x}_0\|} + \lim_{\underline{x} \to \underline{x}_0} \frac{B(\underline{x} - \underline{x}_0)}{\|\underline{x} - \underline{x}_0\|}.$$

$$= \lim_{\underline{x} \to \underline{x}_0} \frac{B(\underline{x} - \underline{x}_0)}{\|\underline{x} - \underline{x}_0\|}.$$

LET $\underline{x} = \underline{x}_0 + t\underline{v}$

$$\Rightarrow \frac{B(t\underline{v})}{\|t\underline{v}\|} = \frac{tB\underline{v}}{|t|\,\|\underline{v}\|} = \pm \frac{B\underline{v}}{\|\underline{v}\|} = 0$$

SO $B\underline{v} = \underline{0}$ ALL $\underline{v}$. THIS IS POSSIBLE ONLY IF B IS THE ZERO MATRIX. $\square$

DEFINITION: LET $\underline{x}_1, \underline{x}_2, \ldots$ BE A SEQUENCE OF VECTORS OF $\mathbb{R}^n$.

WE SAY

$$\lim_{n \to \infty} \underline{x}_n = \underline{x}$$

IF, FOR ANY $\varepsilon > 0$ THERE EXISTS N SO THAT $n > N$ $\Rightarrow$ $\|\underline{x}_n - \underline{x}\| < \varepsilon$.

IF $f: \mathbb{R}^n \to \mathbb{R}^n$, A SEQUENCE OF VECTORS APPROXIMATING $\underline{x}^*$, WITH $f(\underline{x}^*) = \underline{0}$ IS GIVEN BY AN INITIAL GUESS $\underline{x}_0$, AND UPDATE RULE

$$\underline{x}_{n+1} = \underline{x}_n - \left(f'(\underline{x}_n)\right)^{-1} \cdot f(\underline{x}_n)$$

INVERSE MATRIX.

EXAMPLE: $x^2+y^2=2$
$x^2-y^2=1$

SOLUTION: $2x^2=3$, $y^2=\frac{1}{2}$
$x^2=3/2$.

$$f(x,y)=\begin{pmatrix} x^2+y^2-2 \\ x^2-y^2-1 \end{pmatrix}$$

$$f'(x,y)=\begin{pmatrix} 2x & 2y \\ 2x & -2y \end{pmatrix}$$

$$(f'(x,y))^{-1}=\begin{pmatrix} \frac{1}{4x} & \frac{1}{4x} \\ \frac{1}{4y} & -\frac{1}{4y} \end{pmatrix}$$

UPDATE:

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \frac{1}{4x} & \frac{1}{4x} \\ \frac{1}{4y} & -\frac{1}{4y} \end{pmatrix} \cdot \begin{pmatrix} x^2+y^2-2 \\ x^2-y^2-1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{2x^2+3}{4x} \\ \frac{2y^2+1}{4y} \end{pmatrix}$$

$\underline{x}_0=\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $x_1=\begin{pmatrix} 1.25 \\ .75 \end{pmatrix}$,

$x_2=\begin{pmatrix} 1.225 \\ 1.70833 \end{pmatrix}$, $x_3=\begin{pmatrix} 1.22574 \\ 0.707108 \end{pmatrix}$

SOL'N: $\begin{pmatrix} 1.224744\ldots \\ 0.707107\ldots \end{pmatrix}$