Matrix approach to chain rule:

If
$$\begin{tikzcd} & x \arrow[dl] & r \arrow[l] \arrow[dl] \\ z & & \\ & y \arrow[ul] & \theta \arrow[l] \arrow[ul] \end{tikzcd}$$
& $z(x,y)$, $x(r,\theta)$, & $y(r,\theta)$ are differentiable, then $dz = \begin{bmatrix} \frac{\partial z}{\partial x}, & \frac{\partial z}{\partial y} \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix}$

& $\begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial r}, & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r}, & \frac{\partial y}{\partial \theta} \end{bmatrix} \begin{bmatrix} dr \\ d\theta \end{bmatrix}$. Therefore,

$$dz = \begin{bmatrix} \frac{\partial z}{\partial x}, & \frac{\partial z}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial r}, & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r}, & \frac{\partial y}{\partial \theta} \end{bmatrix} \begin{bmatrix} dr \\ d\theta \end{bmatrix} = \begin{bmatrix} \frac{\partial z}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial r}, & \frac{\partial z}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial \theta} \end{bmatrix} \begin{bmatrix} dr \\ d\theta \end{bmatrix}$$

and $dz = \begin{bmatrix} \frac{\partial z}{\partial r}, & \frac{\partial z}{\partial \theta} \end{bmatrix} \begin{bmatrix} dr \\ d\theta \end{bmatrix}$, so $\begin{cases} \frac{\partial z}{\partial r} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial r} \\ \frac{\partial z}{\partial \theta} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial \theta} \end{cases}$

In matrix form, $\begin{bmatrix} \frac{\partial z}{\partial r}, & \frac{\partial z}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \frac{\partial z}{\partial x}, & \frac{\partial z}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial r}, & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r}, & \frac{\partial y}{\partial \theta} \end{bmatrix}$

If we want to reverse the previous page's computation, then we need to invert a matrix:

$$\left[\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right] = \left[\frac{\partial z}{\partial r}, \frac{\partial z}{\partial \theta}\right]\begin{bmatrix} \frac{\partial x}{\partial r}, & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r}, & \frac{\partial y}{\partial \theta} \end{bmatrix}^{-1}$$

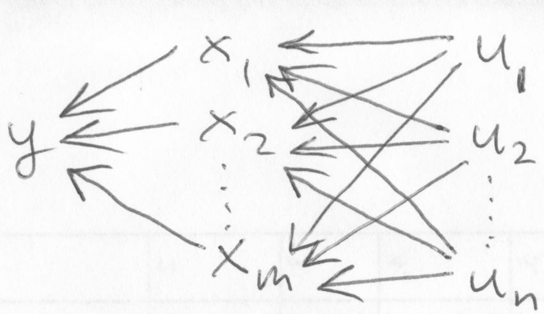
(For 2×2 matrices, $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.)

It is also true that $\left[\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right] = \left[\frac{\partial z}{\partial r}, \frac{\partial z}{\partial \theta}\right]\begin{bmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{bmatrix}$.

Since this is true for all differentiable z,

$$\begin{bmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix}^{-1}$$

General situation:



(Assume functions involved are differentiable.)

$$dy = \begin{bmatrix} \frac{\partial y}{\partial x_1} & \cdots & \frac{\partial y}{\partial x_m} \end{bmatrix} \begin{bmatrix} dx_1 \\ \vdots \\ dx_m \end{bmatrix} \quad \& \quad \begin{bmatrix} dx_1 \\ \vdots \\ dx_m \end{bmatrix} = \begin{bmatrix} \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_1}{\partial u_n} \\ \vdots & & \vdots \\ \frac{\partial x_m}{\partial u_1} & \cdots & \frac{\partial x_m}{\partial u_n} \end{bmatrix} \begin{bmatrix} du_1 \\ \vdots \\ du_n \end{bmatrix}$$

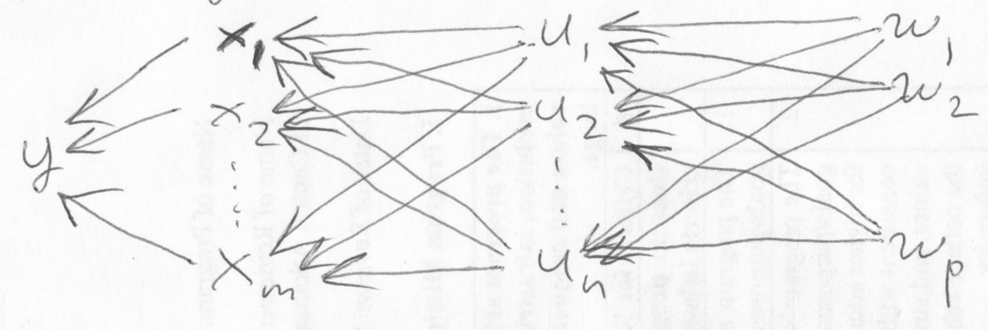
$$dy = \begin{bmatrix} \frac{\partial y}{\partial x_1} & \cdots & \frac{\partial y}{\partial x_m} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_1}{\partial u_n} \\ \vdots & & \vdots \\ \frac{\partial x_m}{\partial u_1} & \cdots & \frac{\partial x_m}{\partial u_n} \end{bmatrix} \begin{bmatrix} du_1 \\ \vdots \\ du_n \end{bmatrix} = \begin{bmatrix} \frac{\partial y}{\partial u_1} & \cdots & \frac{\partial y}{\partial u_n} \end{bmatrix} \begin{bmatrix} du_1 \\ \vdots \\ du_n \end{bmatrix}$$

Therefore, $\frac{\partial y}{\partial u_i} = \sum_{j=1}^{m} \frac{\partial y}{\partial x_j} \frac{\partial x_j}{\partial u_i} = \frac{\partial y}{\partial x_1} \frac{\partial x_1}{\partial u_i} + \cdots + \frac{\partial y}{\partial x_m} \frac{\partial x_m}{\partial u_i}$ because

$$dy = \sum_{i=1}^{m} \frac{\partial y}{\partial x_i} dx_i = \sum_{i=1}^{m} \frac{\partial y}{\partial x_i} \left(\sum_{j=1}^{n} \frac{\partial x_i}{\partial u_j} du_j \right) = \sum_{j=1}^{n} \underbrace{\left(\sum_{i=1}^{m} \frac{\partial y}{\partial x_i} \frac{\partial x_i}{\partial u_j} \right)}_{\partial y / \partial u_j} du_j$$

and $dy = \sum_{j=1}^{n} \frac{\partial y}{\partial u_j} du_j$

A general length-3 "chain":



$$dy = \begin{bmatrix} \frac{\partial y}{\partial x_1} \cdots \frac{\partial y}{\partial x_m} \end{bmatrix} \begin{bmatrix} dx_1 \\ \vdots \\ dx_m \end{bmatrix}$$

$$dx_i = \begin{bmatrix} \frac{\partial x_i}{\partial u_1} \cdots \frac{\partial x_i}{\partial u_n} \end{bmatrix} \begin{bmatrix} du_1 \\ \vdots \\ du_n \end{bmatrix}$$

$$du_j = \begin{bmatrix} \frac{\partial u_j}{\partial w_1} \cdots \frac{\partial u_j}{\partial w_p} \end{bmatrix} \begin{bmatrix} dw_1 \\ \vdots \\ dw_p \end{bmatrix}$$

$$dy = \begin{bmatrix} \frac{\partial y}{\partial x_1} \cdots \frac{\partial y}{\partial x_m} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_1}{\partial u_n} \\ \vdots & & \vdots \\ \frac{\partial x_m}{\partial u_1} & \cdots & \frac{\partial x_m}{\partial u_n} \end{bmatrix} \begin{bmatrix} \frac{\partial u_1}{\partial w_1} & \cdots & \frac{\partial u_1}{\partial w_p} \\ \vdots & & \vdots \\ \frac{\partial u_n}{\partial w_1} & \cdots & \frac{\partial u_n}{\partial w_p} \end{bmatrix} \begin{bmatrix} dw_1 \\ \vdots \\ dw_p \end{bmatrix}$$

Therefore, $\frac{\partial y}{\partial w_k} = \sum_{i=1}^{m} \sum_{j=1}^{n} \left(\frac{\partial x_i}{\partial u_j} \frac{\partial u_j}{\partial w_k} \right)$ because

$$dy = \sum_{i=1}^{m} \frac{\partial y}{\partial x_i} dx_i = \sum_{i=1}^{m} \frac{\partial y}{\partial x_i} \left(\sum_{j=1}^{n} \frac{\partial x_i}{\partial u_j} du_j \right) = \sum_{i=1}^{m} \frac{\partial y}{\partial x_i} \left(\sum_{j=1}^{n} \frac{\partial x_i}{\partial u_j} \left(\sum_{k=1}^{p} \frac{\partial u_j}{\partial w_k} dw_k \right) \right)$$

$$= \sum_{k=1}^{p} \underbrace{\left(\sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\partial x_i}{\partial u_j} \frac{\partial u_j}{\partial w_k} \right)}_{\partial y / \partial w_k} dw_k \quad \& \quad \sum_{k=1}^{p} \frac{\partial y}{\partial w_k} dw_k = dy$$