eq 1.1

(D,) (N,D)

$$d\beta = \frac{\partial out}{\partial \beta} \cdot \sum_{i=1}^{N} dout$$

$$= \sum_{i=1}^{N} dout$$

eq 1.2

(D,) (N,D)

$$d\gamma = \frac{\partial \gamma \hat{x}}{\partial \gamma} \cdot \sum_{i=1}^{N} dout$$

$$= \sum_{i=1}^{N} dout \cdot \hat{x}_i$$

now, to find dx

eq 1.3

(N,D) (N,D)

$$d\hat{x} = \frac{\partial \gamma \hat{x}}{\partial \hat{x}} \cdot dout = \gamma \cdot dout$$

eq 1.4

Let's assume $a = x - \mu$, and $b = \frac{1}{\sqrt{\sigma^2 + \epsilon}}$

$$\underset{(N,D)}{d(1\_4)} = d\hat{x} \cdot \frac{\partial \hat{x}}{\partial a} = d\hat{x} \cdot \frac{\partial ab}{\partial a} = d\hat{x} \cdot b = \frac{\overset{(N,D)}{d\hat{x}}}{\sqrt{\sigma^2+\epsilon}}$$

eq 1.5

$$\underset{(D,\ )}{d(1\_5)} = \sum_{i=1}^{N} d\hat{x} \cdot \frac{\partial \hat{x}}{\partial b} = \sum_{i=1}^{N} d\hat{x} \cdot a = \overbrace{\sum_{i=1}^{N} d\hat{x} \cdot \overbrace{(x_i - \mu)}^{(N,D)}}^{(D,\ )}$$

eq 1.51

$$\frac{d\left(\frac{1}{x}\right)}{dx} = -\frac{1}{x^2}$$

therefore, $\underset{(D,\ )}{d(1\_51)} = -b^2 \cdot d(1\_5) = \frac{\overset{(D,\ )}{-d(1\_5)}}{\sigma^2+\epsilon}$

eq 1.52

$$\frac{d\left(\sqrt{x+\epsilon}\right)}{dx} = \frac{d\left((x+\epsilon)^{\frac{1}{2}}\right)}{dx} = \frac{1}{2}(x+\epsilon)^{-\frac{1}{2}}$$

$$= \frac{1}{2} \frac{1}{\sqrt{x+\epsilon}}$$

therefore, $d(1\_52) = \frac{1}{2} \cdot \frac{1}{\sqrt{\sigma^2+\epsilon}} \cdot d(1\_51)$

eq 1.53

$$\frac{\partial\left(\frac{1}{N}\sum_{i=1}^{N} x\right)}{\partial x} = \frac{1}{N}$$

, but since $d(1\text{-}53)$ is of size (N,D)

then

$$\overset{(N,D)}{d(1\text{-}53)} = \frac{1}{N}\, \overbrace{\mathbb{1}^{N\times D} \times \overset{(D,\ )}{d(1\text{-}52)}}^{(N,D)}$$

eq 1.54

$$\frac{d\, x^2}{x} = 2x$$

$$\therefore\ d(1\text{-}54) = 2a \cdot d(1\text{-}53)$$

$$= 2(x-\mu) \cdot d(1\text{-}53)$$

eq 1.6

$$\underset{}{\overset{(D,1)}{d\mu}} = -\sum_{i=1}^{N} \overbrace{\left[\overset{(N,D)}{d(1\text{-}4)} + \overset{(N,D)}{d(1\text{-}54)}\right]}^{(D,\ )}$$

(since $\frac{d\, x-\mu}{d\mu} = -1$)

eq 1.7

$$\overset{(N,D)}{d(1.7)} = \overset{(N,D)}{d(1-4)} + \overset{(N,D)}{d(1-54)}$$

$$\left(\text{since } \frac{d\; x-N}{dx} = 1\right)$$

eq 1.8

$$\overset{(N\times D)}{d(1-8)} = \overbrace{\frac{1}{N} \cdot 1^{N\times D} \times \overset{(D,1)}{d\,mu}}^{(N\times D)}$$

$$\ldots \text{finally, } \overset{(N,D)}{dx} = \overset{(N,D)}{d(1-7)} + \overset{(N,D)}{d(1-8)}$$