ECE 792 (061) HW 4 SOMSHUBHRA ROY

1) Given a Normalised Flow has 1 affine coupling

$f : x \mapsto z$

$x, z \in \mathbb{R}^D$

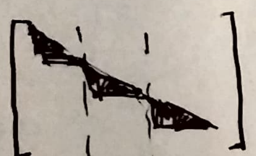
ie $x = \begin{bmatrix} x_1 \\ \vdots \\ x_d \\ \vdots \\ x_D \end{bmatrix}$ $z = \begin{bmatrix} z_1 \\ \vdots \\ z_d \\ \vdots \\ z_D \end{bmatrix}$

There is a modification in during sampling / transformation of x

instead of $z_{1:d} = x_{1:d}$

Let $z_{1:d} = A x_{1:d}$.

where A is a block Lower Triangular Matrix. such that

$A =$  $d \times d$.

m_1 m_1+m_2

where the triangles are the only non-zero elements of the matrix.

such that

$z_1 = A_{11} x_1$

$z_2 = A_{21} x_1 + A_{22} x_2$

$\vdots$

$z_{m_1} = A_{m_1 1} x_1 + A_{m_1 2} x_2 + \dots + A_{m_1 m_1} x_{m_1}$

$z_{m_1+1} = A_{m_1+1\, m_1+1} x_{m_1+1}$

$\vdots$

$z_{m_1+m_2} = A_{m_1+m_2,\, m_1+1} x_{m_1+1} + A_{m_1+m_2,\, m_1+2} x_{m_1+2} + \dots A_{m_1+m_2,\, m_1+m_2} x_{m_1+m_2}$

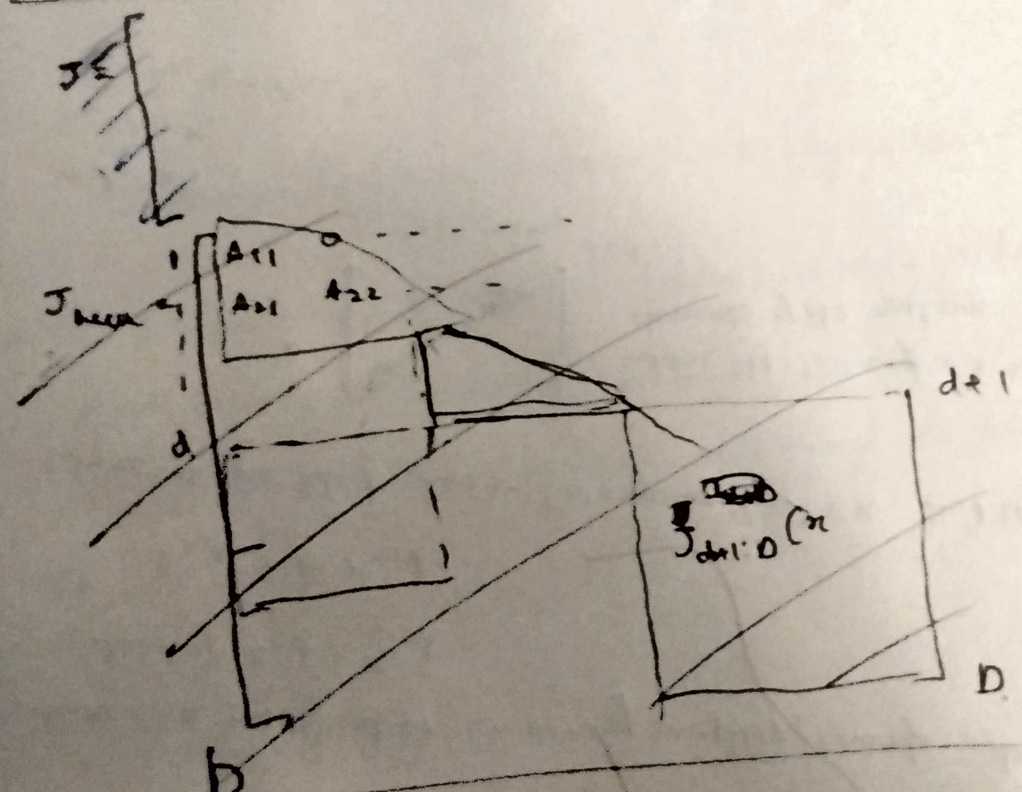
$\vdots$

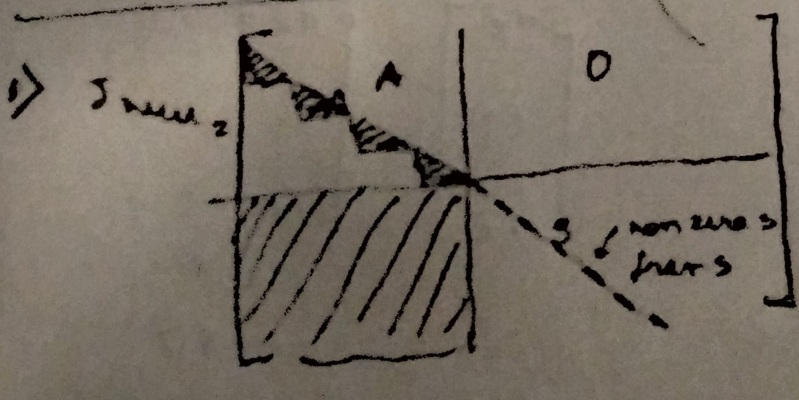
$z_d = A_{d,\, m_1+\dots m_{n-1}+1} x_{m_1+\dots+m_{n-1}+1} + \dots\dots$ Add x_d.

where $0 \leq m_1, m_2, \dots m_n \leq d$ & $\sum_{i=1}^{n} m_i = d$.

finally ~~[illegible]~~ $z_{d+1:D} = x_{d+1:D} \odot e^{s(x_{1:d})} + t(x_{1:d})$ where $s(\cdot), t(\cdot)$ map $\mathbb{R}^d \mapsto \mathbb{R}^{D-d}$, & $\odot$ is the elementwise product

$$\therefore \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{m_1} \\ z_{m_1+1} \\ \vdots \\ z_{m_1+m_2} \\ z_{m_1+m_2+1} \\ \vdots \\ z_d \end{bmatrix} = \begin{bmatrix} A_{11} & 0 & \cdots & & & \\ A_{21} & A_{22} & 0 & \cdots & & \\ A_{m_1 1} & A_{m_1 2} & \cdots & A_{m_1 m_1} & & \\ 0 & 0 & \cdots & 0 & A_{m_1+1\, m_1+1} & 0 \cdots \\ & & & & \vdots & \\ & & & & A_{m_1+m_2\, m_2 1} & \cdots \\ & & & & & A \\ & & & & & \vdots \\ & & & & & A_{d m_1} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix}$$





$$s = e^{s(x_{1:d})}$$

2) ~~The deter~~

$\det |J| = \prod \det(\text{triangular blocks of } A) \odot \text{diag}(S)$

Increasing the number of triangular blocks allows the model to capture more complex correlation in data allows the model to be more expressive, However it also increase the number of learnable parameters thus increasing computational complexity & overfitting.

3) The modified Jacobian changes the geometry of the transformed data manifold & captures more complex data correlations which increases the expressiveness of the model thus increasing the richness of the synthesised data.