Generative Adversarial Networks

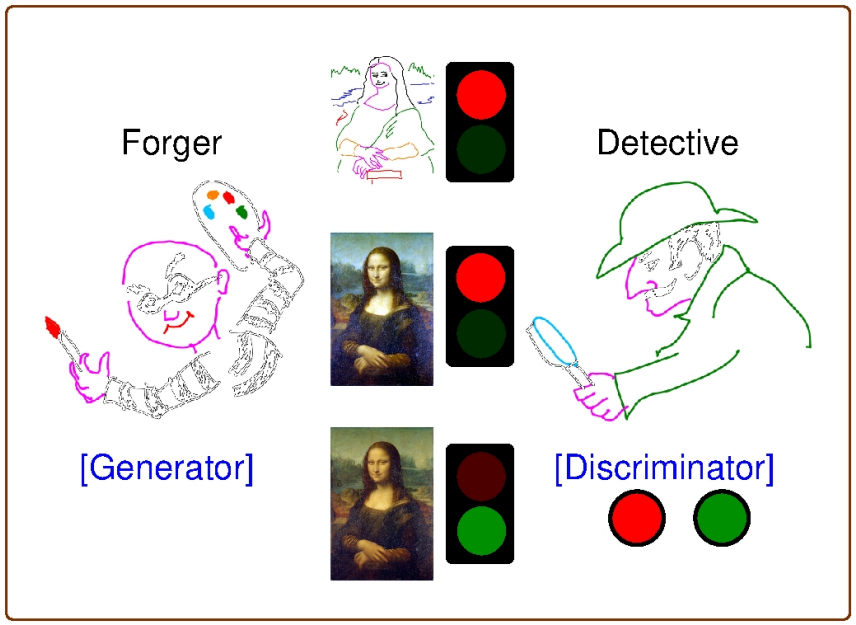
Sumantra Dutta Roy
Department of Electrical Engineering
Indian Institute of Technology Delhi
Hauz Khas, New Delhi - 110 016, INDIA.
http://www.cse.iitd.ac.in/∼sumantra
sumantra@ee.iitd.ac.in

The basic GAN concept

https://upload.wikimedia.org/wikipedia/commons/f/fe/Ian_Goodfellow.jpg

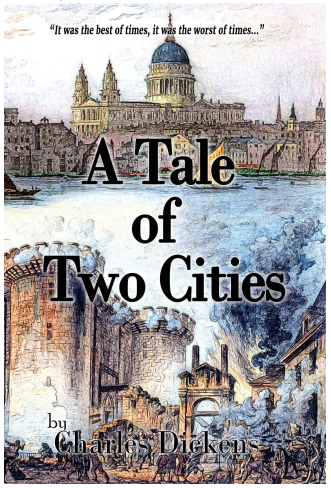
- **Ian J. Goodfellow [1987-]**
- [I. J. Goodfellow *et al.*, NIPS'14]
- [J. Schmidhuber'91], [O. Neimitalo'10], [Li, Gauci, Gross'13]

https://upload.wikimedia.org/wikipedia/commons/e/ec/Mona_Lisa%2C_by_Leonardo_da_Vinci%2C_from_C2RMF_retouched.jpg



A Tale of Two (Atro)Cities

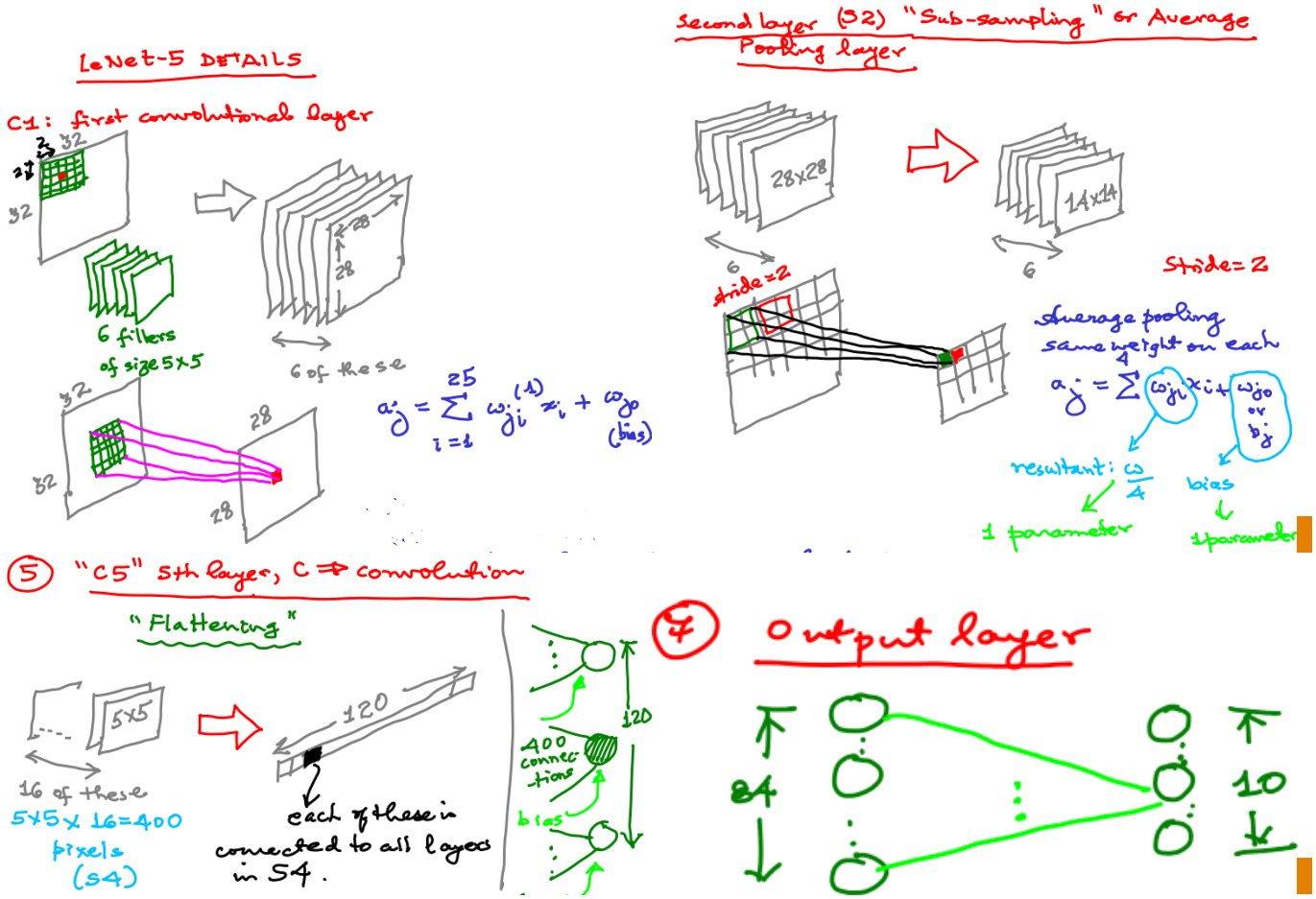
https://m.media-amazon.com/images/I/91-J0ie5SmL._SL1500_.jpg



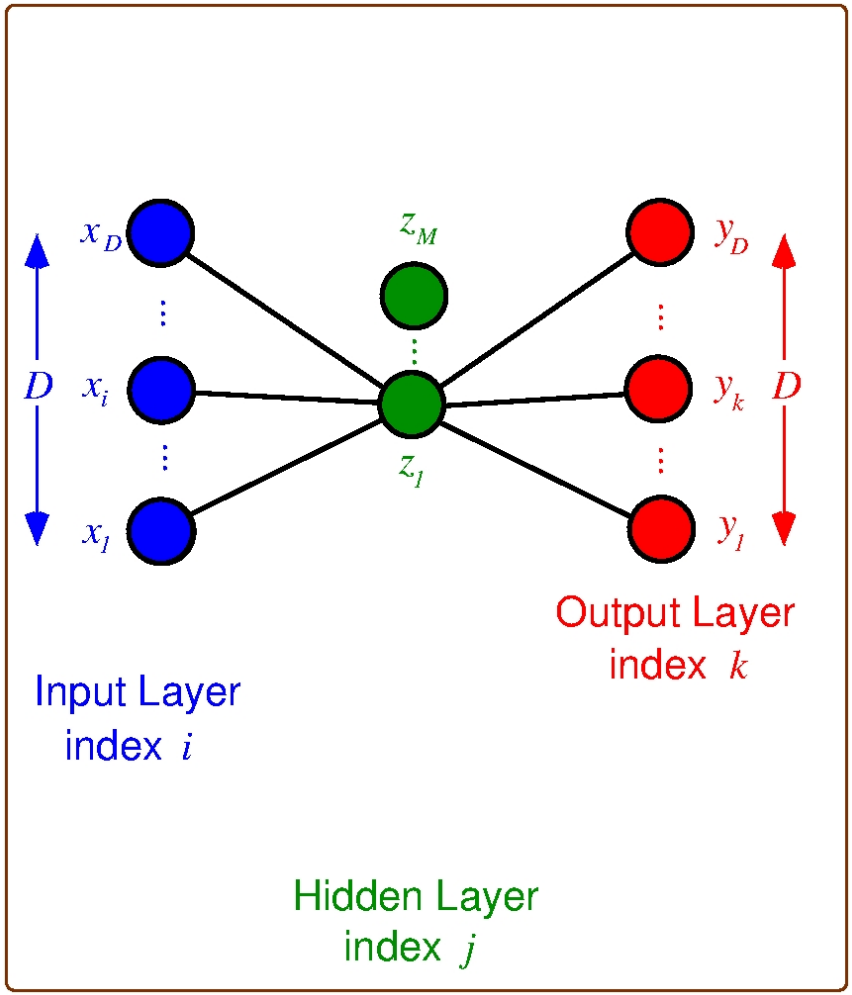
- Generative Models: sampling: new point in input space (∼GMM)
- Gen: new samples, Disc: real/fake
- 0-Sum Adv game, till Disc fooled 50%
- Disc can be discarded post-training
- GANs use both bow-tie & contractive

- Bow-tie: o/p, i/p same dim: Generator
- Simple Example: Auto-Encoder
- Contractive: 2-class classifier: Discriminator
- Simple Example: Convolutional NN (Oxymoron)
- Neither 'convolution', nor 'neural network' [00:48]

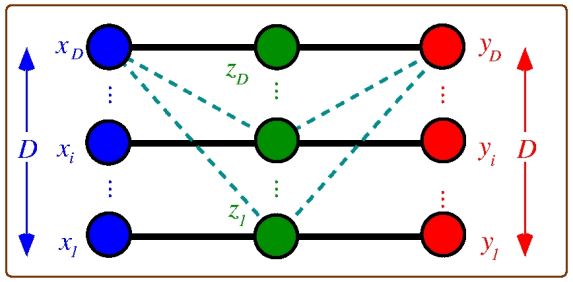
Contractive: Discrimininator

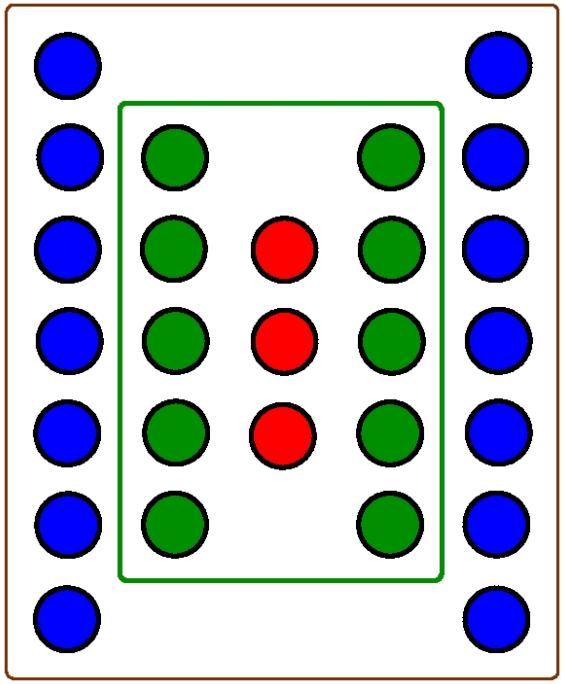


Bow-tie: Autoencoder

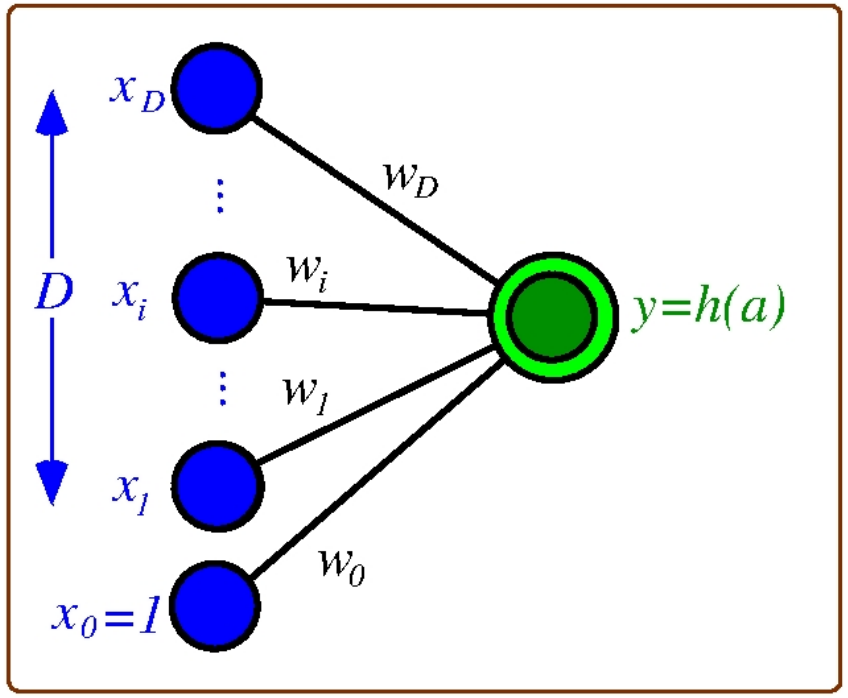


Bow-tie: Autoencoder (contd)

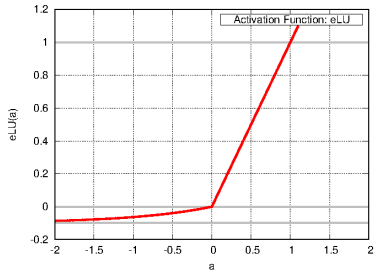




Serrano's Perceptron-like NNs



- A single-layer NN
- #neurons?[04:00]
- $h(a)$? $\sigma(a)$
- $y_i = \mathbf{w}^T \mathbf{x}_i + borw_0$
- b/w_0: bias/thresh

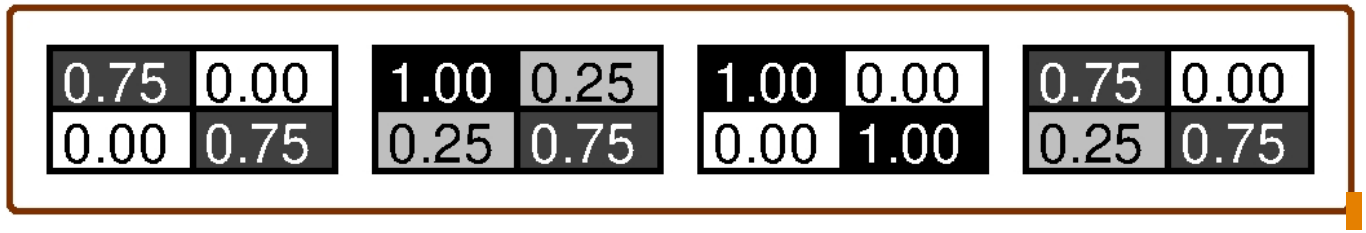


- Activation fn: σ, $tanh$, RelU, Leaky ReLU, eLU
- $LeakyRelU(a,\alpha) \stackrel{\triangle}{=} \max(\alpha a, a),\ \alpha \in (0,1)$
- $eLU(a,\alpha) \stackrel{\triangle}{=} \left\{ \begin{array}{c} a, a > 0) \\ \alpha(e^a - 1), a \leq 0 \end{array} \right\}$ [00:50]

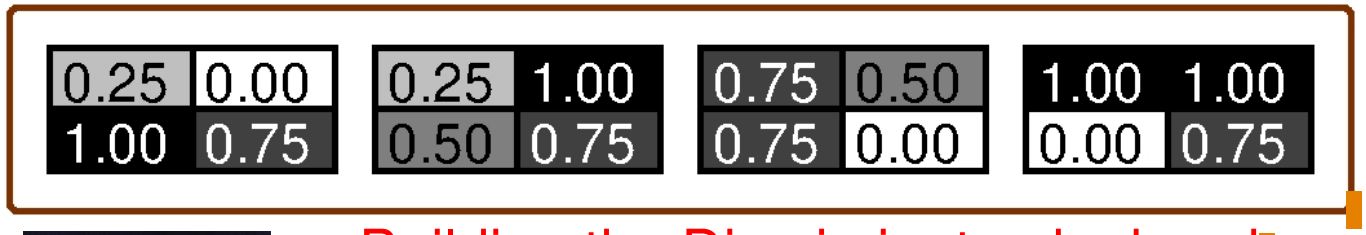
Angels & Demons: \s & Noise

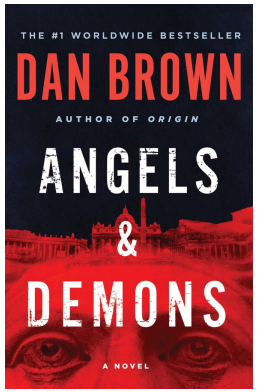
https://d28hgpri8am2if.cloudfront.net/book_images/onix/cvr9780743493468/angels-demons-9780743493468_xlg.jpg

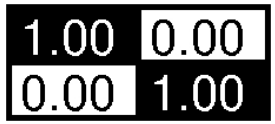
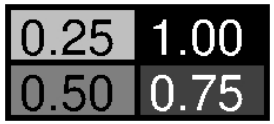
Acceptable backslashes (2×2 images):



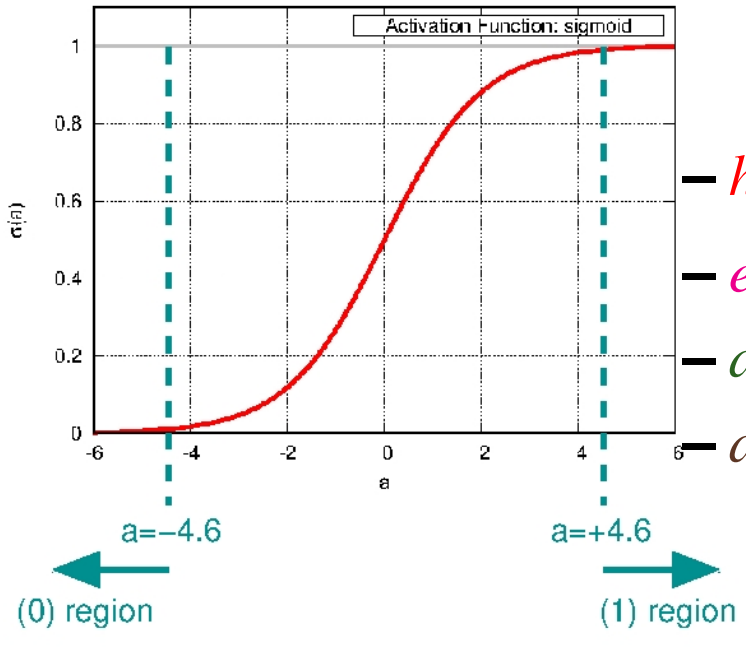
Noise:





- Building the Discriminator, by hand
- 1 possible fn: $\sum$ Main Diag - $\sum$ Other
-  $= +2.00$  $= -0.50$
- Backslash: high response, noise: low

- To translate the fn into a NN: Σ weights +1, -1
- To translate the score $\in (-\infty, +\infty)$ into a prob:



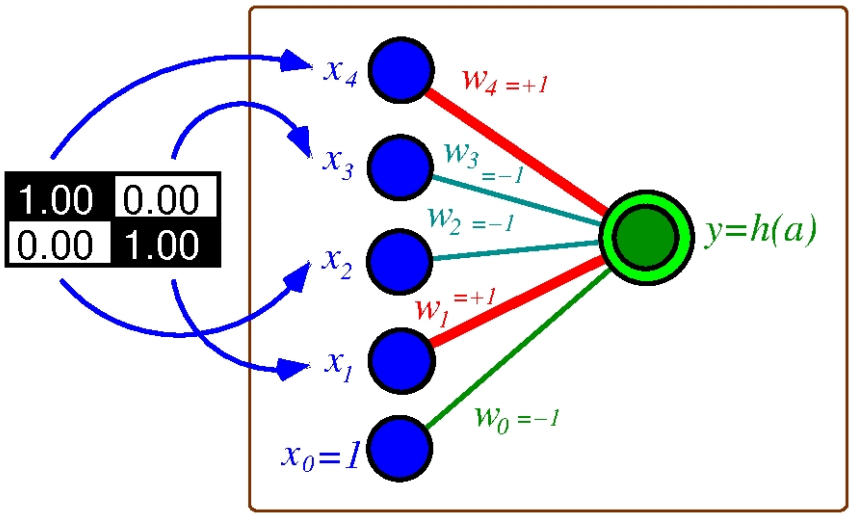
$h(a) = \frac{1}{1+e^{-a}}$

$e^{4.6} \approx 100$

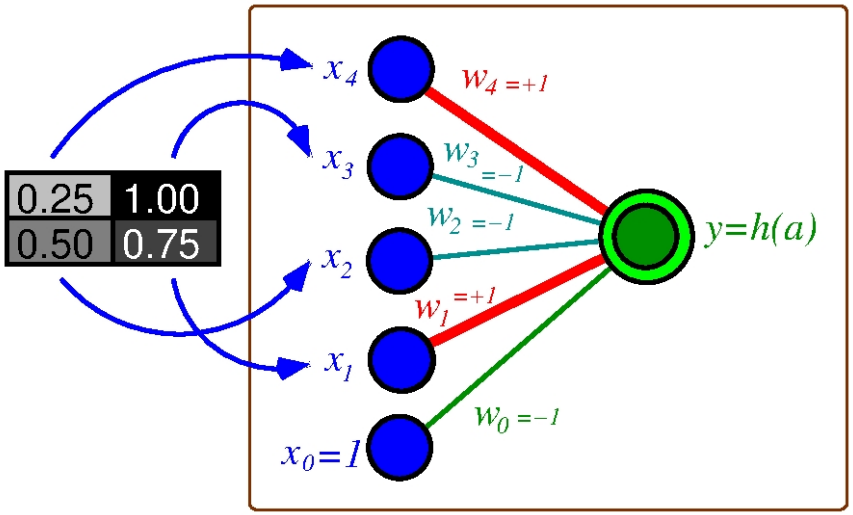
$a = -4.6 : h(a) \approx 0.01$

$a = +4.6 : h(a) \approx 0.99$

Discriminator: Backslash & Noise



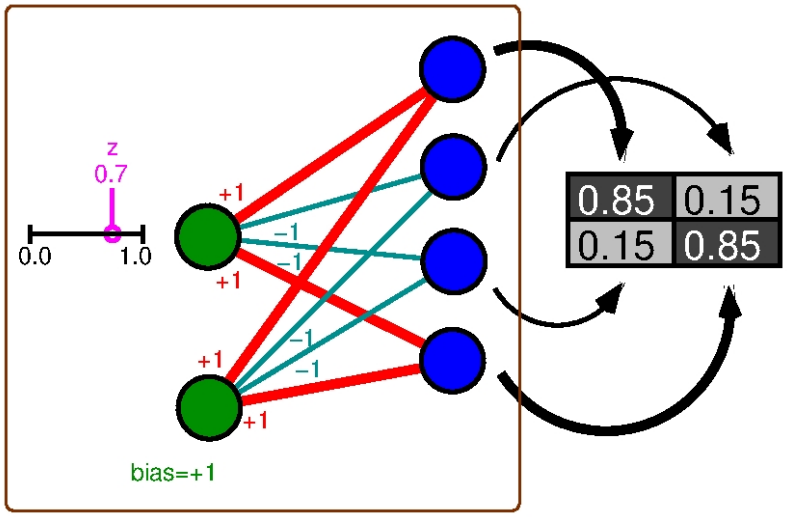
- $a = 1.00(+1) + 0.00(-1) + 0.00(-1) + 1.00(+1) + 1(-1) = +1.00$
- $\sigma(+1.00) = 0.73$
- $\sigma(\cdot) : \uparrow \# \rightarrow \uparrow prob$



- $a = 0.25(+1) + 1.00(-1) + 0.50(-1) + 0.75(+1) + 1(-1) = -0.50$
- $\sigma(-1.50) = 0.37$
- $\sigma(\cdot) : \downarrow \# \rightarrow \downarrow prob$

Building the Generator

- High (+1) wt: main diag (#1, #4), Low (-1): other



$$\rightarrow a_4 = 0.70(+1) + 1.00(+1) = +1.7$$

$$\rightarrow a_3 = 0.70(-1) + 1.00(-1) = -1.7$$

$$\rightarrow a_2 = 0.70(-1) + 1.00(-1) = -1.7$$

$$\rightarrow a_1 = 0.70(+1) + 1.00(+1) = +1.7$$

- $$\left\{ \begin{array}{l} \sigma(+1.7) = 0.85, \;\; \sigma(-1.7) = 0.15 \\ \sigma(-1.7) = 0.15, \;\; \sigma(+1.7) = 0.85 \end{array} \right\}$$