

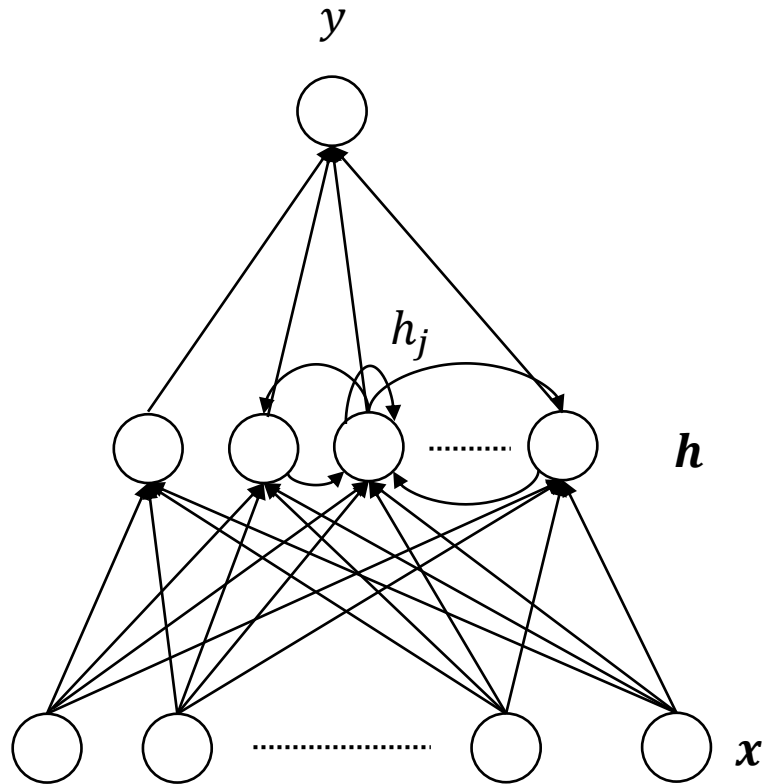
# Machine Learning

# Contents

1. Introduction
2. K-Nearest Neighbor Algorithm
3. LDA(Linear Discriminant Analysis)
4. Perceptron
5. Feed-Forward Neural Networks
- 6. RNN(Recurrent Neural Networks)**
7. SVM(Support Vector Machine)
8. Ensemble Learning
9. CNN(Convolutional Neural Network)
10. PCA(Principal Component Analysis)
11. ICA(Independent Component Analysis)
12. Clustering
13. GAN(Generative Adversarial Network)

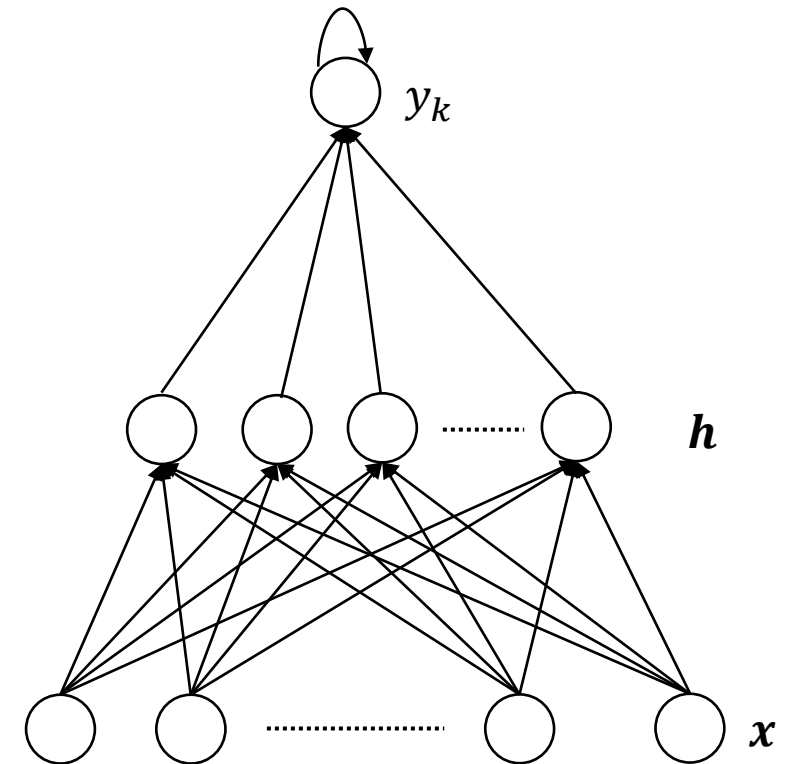
# 6.1. Recurrent Neural Networks

Universal approximation theorem for time series data



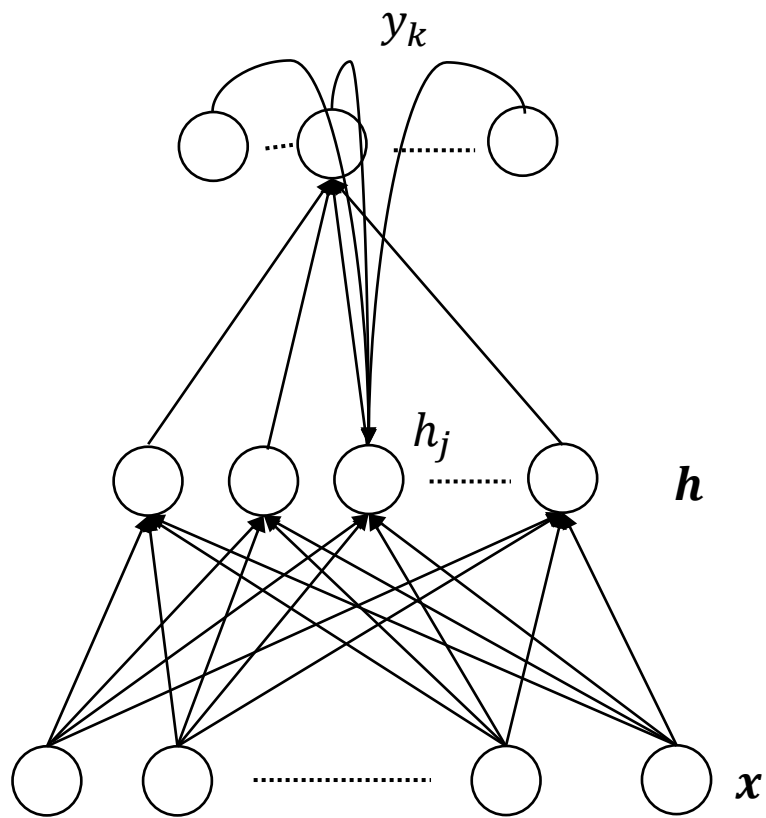
Recurrent connection for hidden nodes

$$\hat{h}_j(t) = \sum_i w_{ji} x_i + \sum_k T_{jk} h_k(t-1) \quad (6.1.1)$$



Recurrent connection for output nodes

$$\hat{y}_k(t) = \sum_j v_{kj} h_j + T_{kk} y_k(t-1) \quad (6.1.2)$$



Recurrent connection from output to hidden nodes

$$\hat{h}_j(t) = \sum_i w_{ji} x_i + \sum_k T_{jk} y_k(t-1) \quad (6.1.3)$$

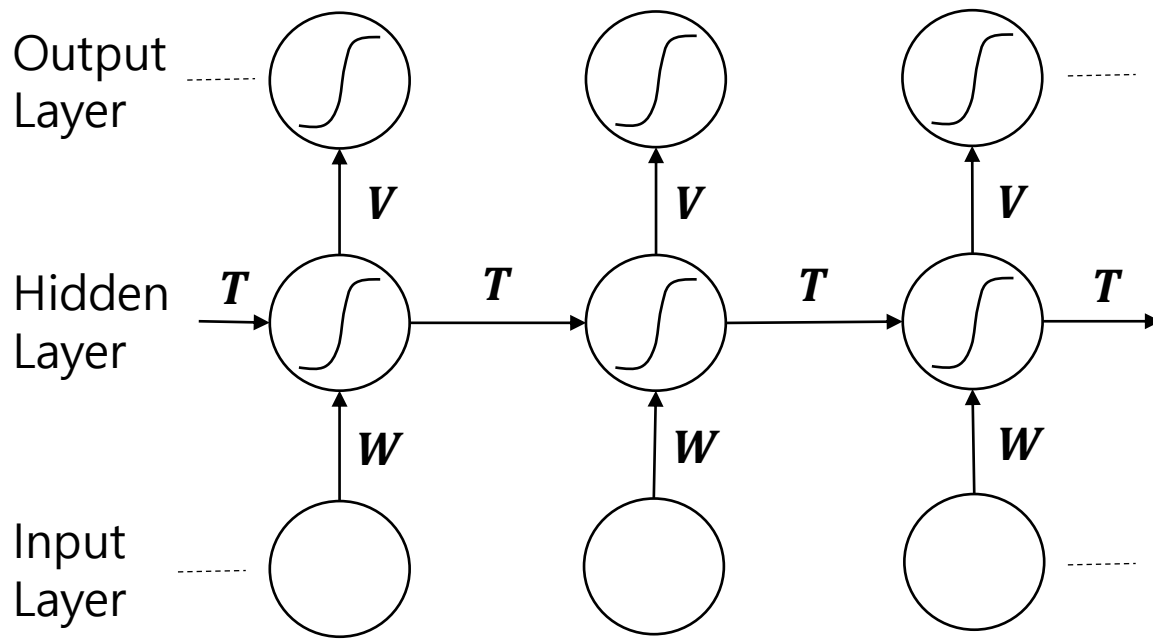
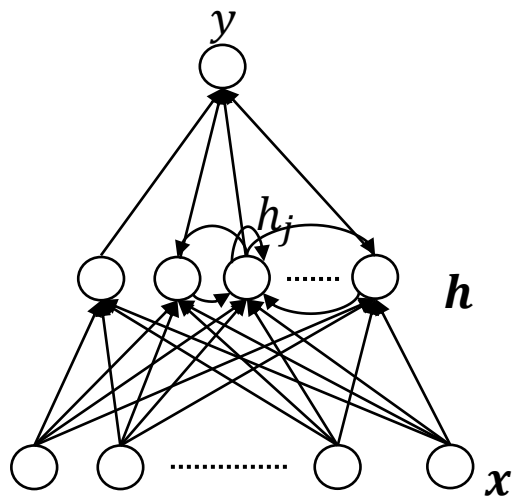


그림 6.4. 시간에 대하여 펼쳐진 은닉노드 회귀연결 신경회로망. 각 원은 특정 시간의 해당 층 노드들을 뜻함.  $W$ 는 입력층에서 은닉층으로의 연결 가중치,  $V$ 는 은닉층에서 출력층으로의 연결 가중치,  $T$ 는 은닉층 내부의 회귀 연결 가중치를 각각 뜻함.

$$\delta_j^{(hidden)}(t) = f'(h_j^{(t)}) \left( \sum_{k=1}^M v_{kj} \delta_k + \sum_{k=1}^H T_{jk} \delta_k^{(hidden)}(t+1) \right) \quad (6.1.4)$$

## 6.2. Long Short-Term Memory(LSTM)

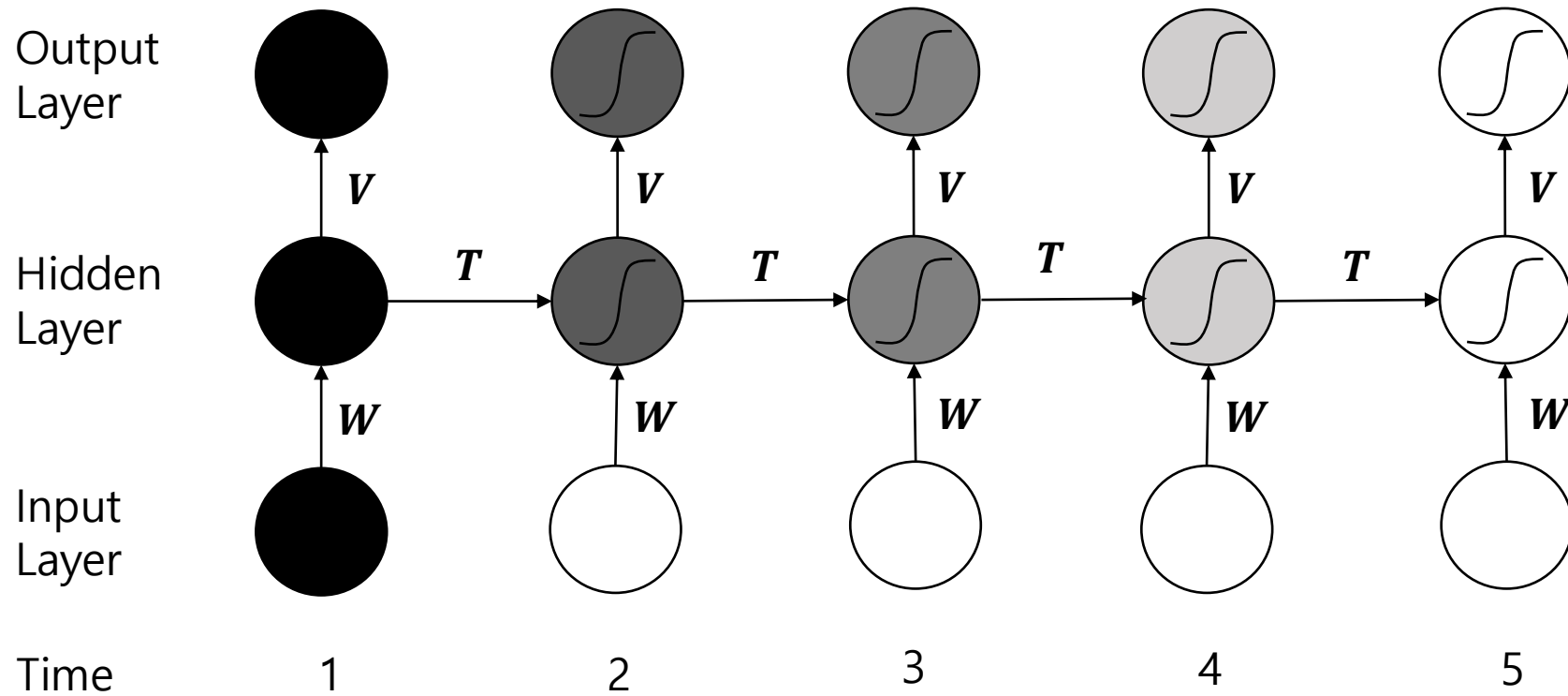


그림 6.5. 회귀신경회로망에서 시간에 따른 영향의 감소, 색깔의 강도는 민감도의 강도를 나타냄.

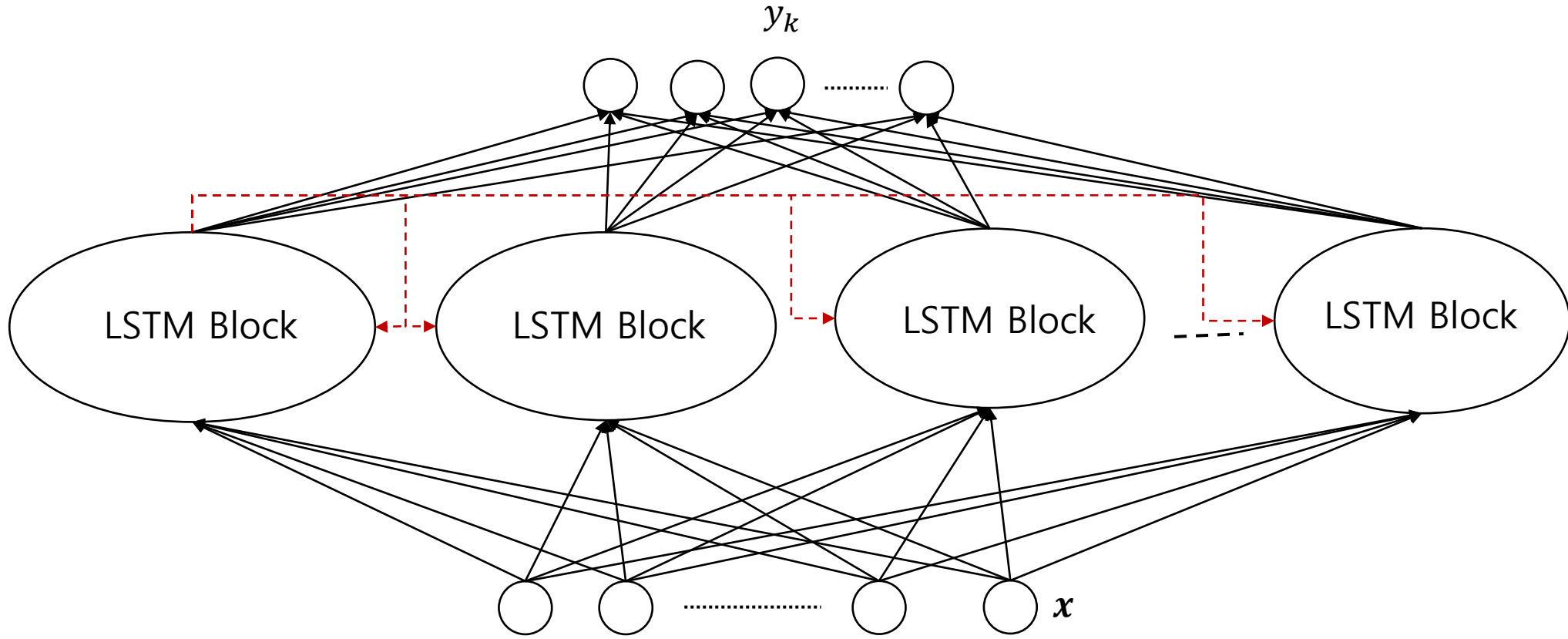
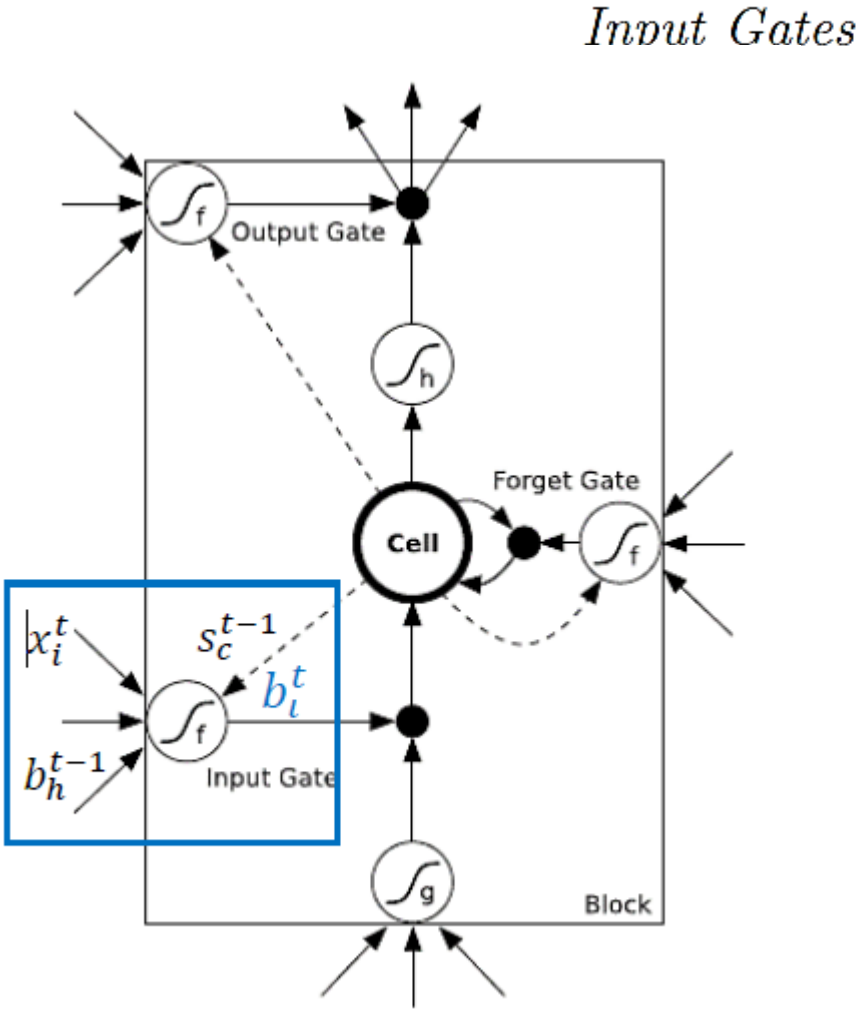


그림 6.6. LSTM 블록을 지닌 회귀 신경화로망. LSTM 블록 사이의 연결은 한 방향만을 표현하였지만, 모든 LSTM 블록들은 서로 연결되어 있음.





# 6.3. Forward Pass

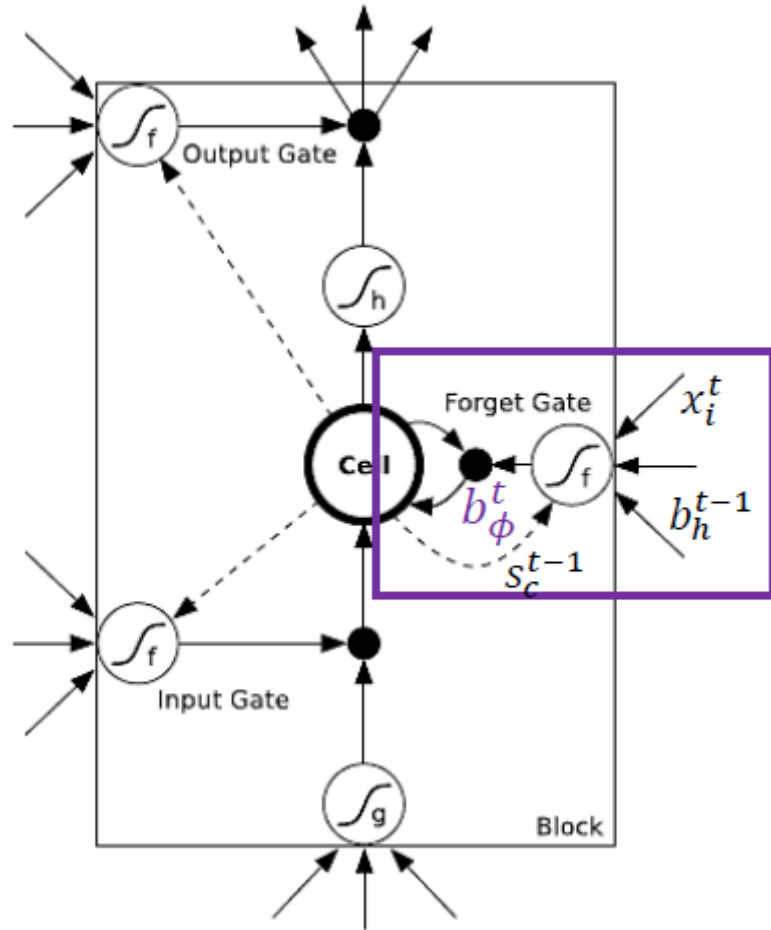


$$a_l^t = \sum_{i=1}^I w_{li} x_i^t + \sum_{h=1}^H w_{lh} b_h^{t-1} + \sum_{c=1}^C w_{lc} s_c^{t-1} \tag{6.3.1}$$

$$b_l^t = f(a_l^t)$$

- I : # inputs
- K : # outputs
- H : # cells in the hidden layer
- C : # cells per a memory
- G : total # inputs to the hidden layer (= 4H)

### Forget Gates



$$a_\phi^t = \sum_{i=1}^I w_{i\phi} x_i^t + \sum_{h=1}^H w_{h\phi} b_h^{t-1} + \sum_{c=1}^C w_{c\phi} s_c^{t-1} \quad (6.3.3)$$

$$b_\phi^t = f(a_\phi^t)$$

I : # inputs

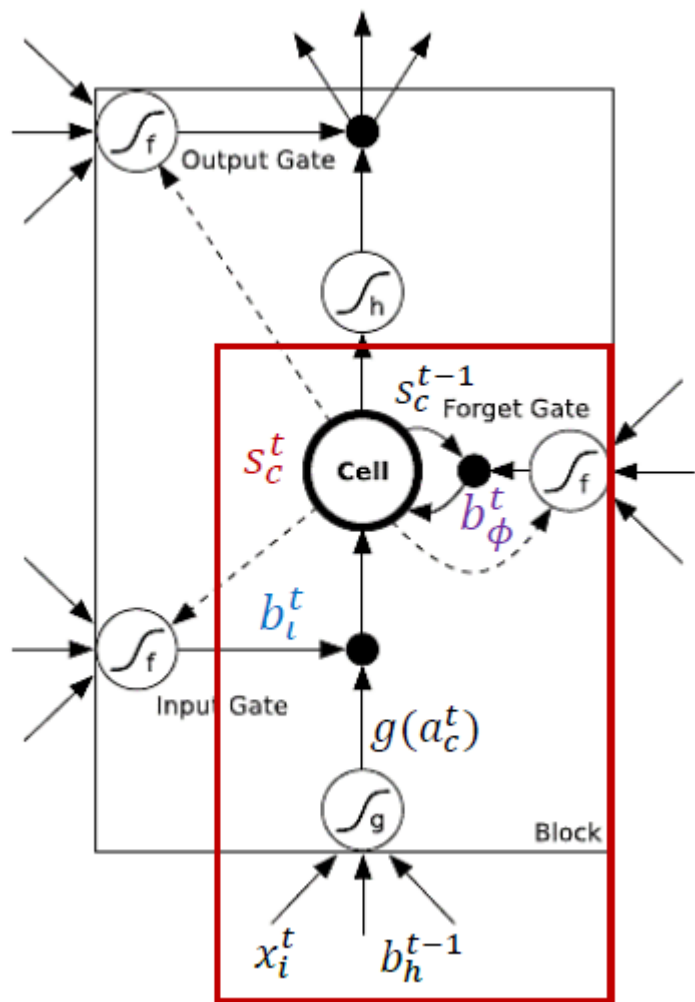
K : # outputs

H : # cells in the hidden layer

C : # cells per a memory

G : total # inputs to the hidden layer (= 4H)

## Cells



$$a_c^t = \sum_{i=1}^I w_{ic} x_i^t + \sum_{h=1}^H w_{hc} b_h^{t-1} \quad (6.3.5)$$

$$s_c^t = b_\phi^t s_c^{t-1} + b_\iota^t g(a_c^t) \quad (6.3.6)$$

I : # inputs

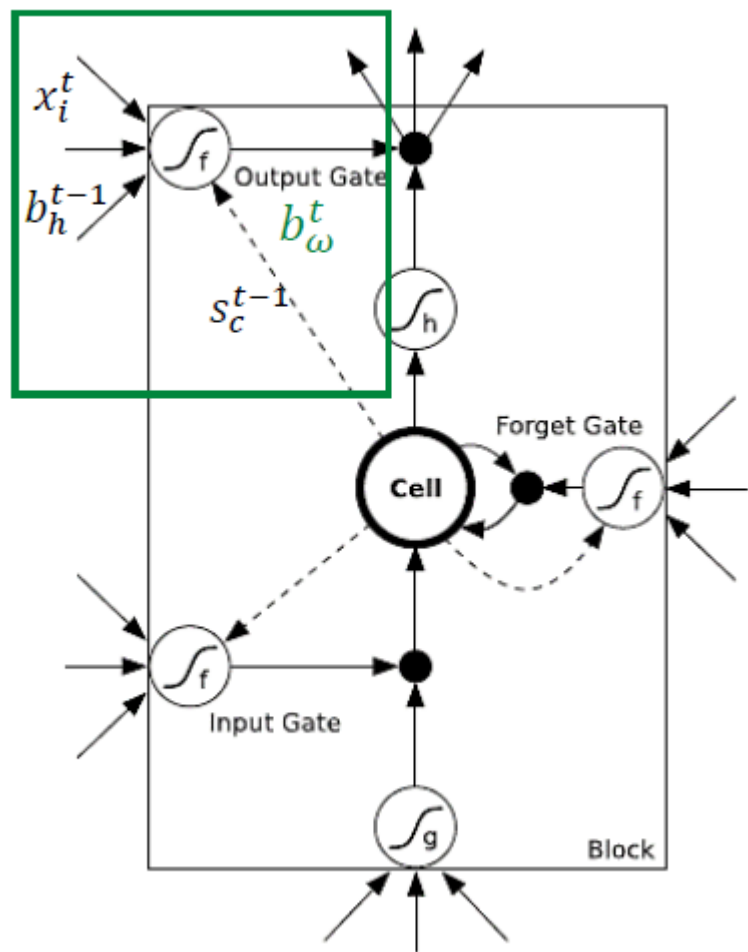
K : # outputs

H : # cells in the hidden layer

C : # cells per a memory

G : total # inputs to the hidden layer (= 4H)

## Output Gates



$$a_\omega^t = \sum_{i=1}^I w_{i\omega} x_i^t + \sum_{h=1}^H w_{h\omega} b_h^{t-1} + \sum_{c=1}^C w_{c\omega} s_c^t \quad (6.3.7)$$

$$b_\omega^t = f(a_\omega^t)$$

I : # inputs

K : # outputs

H : # cells in the hidden layer

C : # cells per a memory

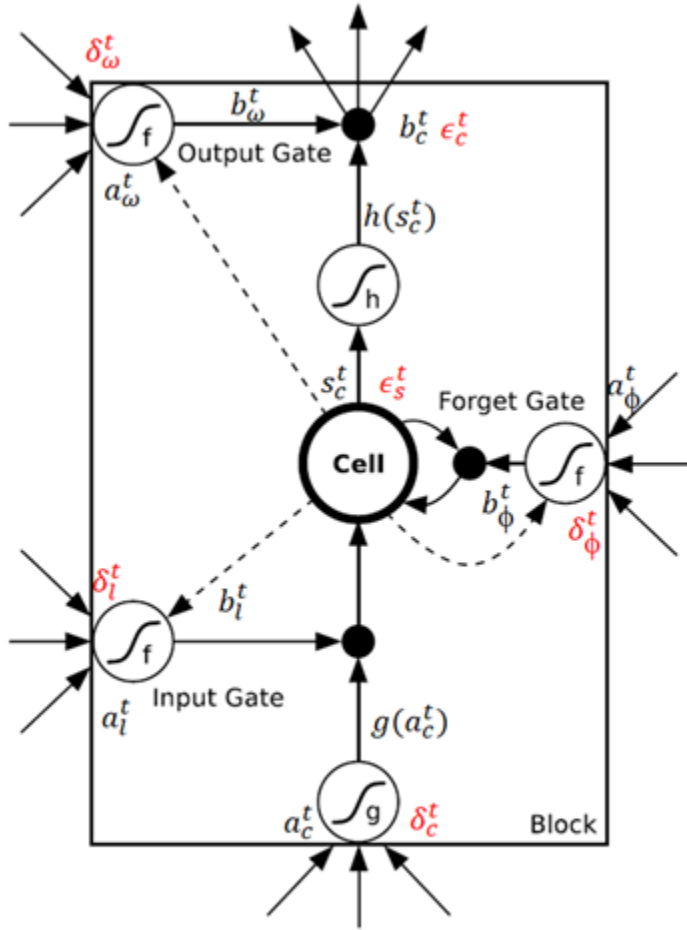
G : total # inputs to the hidden layer (= 4H)



# 6.4. Backward Pass

Loss function:  $L$

Cell Outputs



$$\epsilon_c^t \equiv \frac{\partial L}{\partial b_c^t} = \sum_{k=1}^K w_{ck} \delta_k^t + \sum_{g=1}^G w_{cg} \delta_g^{t+1} \quad (6.4.1)$$

$$a_i^t = \sum_{i=1}^I w_{ii} x_i^t + \sum_{h=1}^H w_{hi} b_h^{t-1} + \sum_{c=1}^C w_{ci} s_c^{t-1} \quad (6.3.1)$$

$$a_\phi^t = \sum_{i=1}^I w_{i\phi} x_i^t + \sum_{h=1}^H w_{h\phi} b_h^{t-1} + \sum_{c=1}^C w_{c\phi} s_c^{t-1} \quad (6.3.3)$$

$$a_c^t = \sum_{i=1}^I w_{ic} x_i^t + \sum_{h=1}^H w_{hc} b_h^{t-1} \quad (6.3.5)$$

$$a_\omega^t = \sum_{i=1}^I w_{i\omega} x_i^t + \sum_{h=1}^H w_{h\omega} b_h^{t-1} + \sum_{c=1}^C w_{c\omega} s_c^t \quad (6.3.7)$$

그림 6.12: 전방향 및 역방향 변수

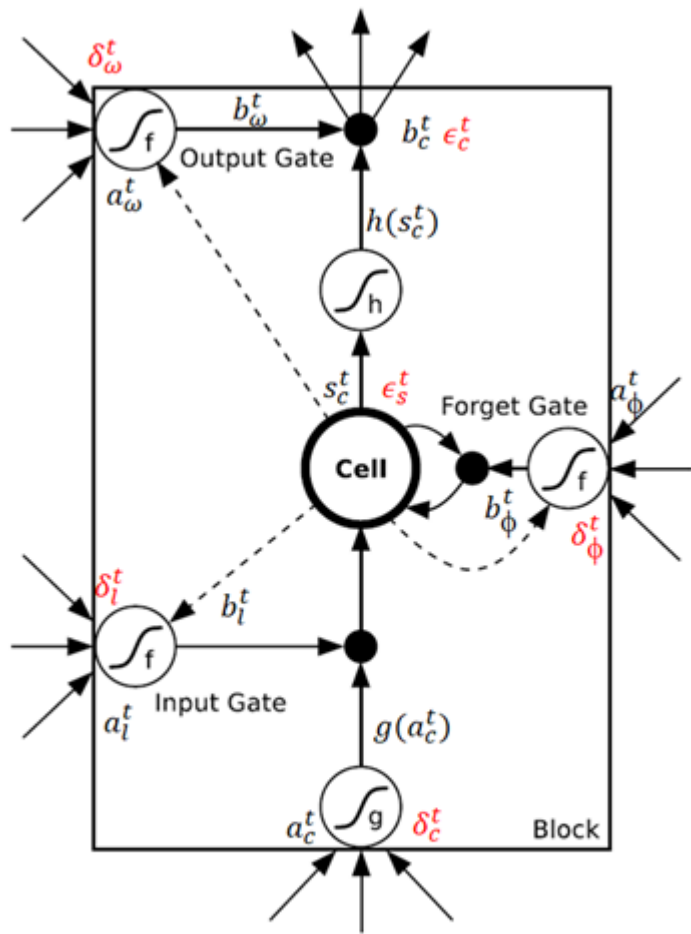


그림 6.12: 전방향 및 역방향 변수

### Output Gates

$$\delta_w^t \equiv \frac{\partial L}{\partial a_w^t} = f'(a_w^t) \sum_{c=1}^C h(s_c^t) \epsilon_c^t \quad (6.4.2)$$

### States

$$\epsilon_s^t \equiv \frac{\partial L}{\partial s_c^t} = b_w^t h'(s_c^t) \epsilon_c^t + \boxed{b_\phi^{t+1} \epsilon_s^{t+1}} + \boxed{w_{cl} \delta_l^{t+1}} + \boxed{w_{c\phi} \delta_\phi^{t+1}} + \boxed{w_{cw} \delta_w^t} \quad (6.4.3)$$

$$s_c^t = \boxed{b_\phi^t s_c^{t-1}} + b_l^t g(a_c^t) \quad (6.3.6)$$

$$a_l^t = \sum_{i=1}^I w_{li} x_i^t + \sum_{h=1}^H w_{lh} b_h^{t-1} + \boxed{\sum_{c=1}^C w_{cl} s_c^{t-1}} \quad (6.3.1)$$

$$a_\phi^t = \sum_{i=1}^I w_{\phi i} x_i^t + \sum_{h=1}^H w_{\phi h} b_h^{t-1} + \boxed{\sum_{c=1}^C w_{c\phi} s_c^{t-1}} \quad (6.3.3)$$

$$a_w^t = \sum_{i=1}^I w_{wi} x_i^t + \sum_{h=1}^H w_{hw} b_h^{t-1} + \boxed{\sum_{c=1}^C w_{cw} s_c^t} \quad (6.3.7)$$

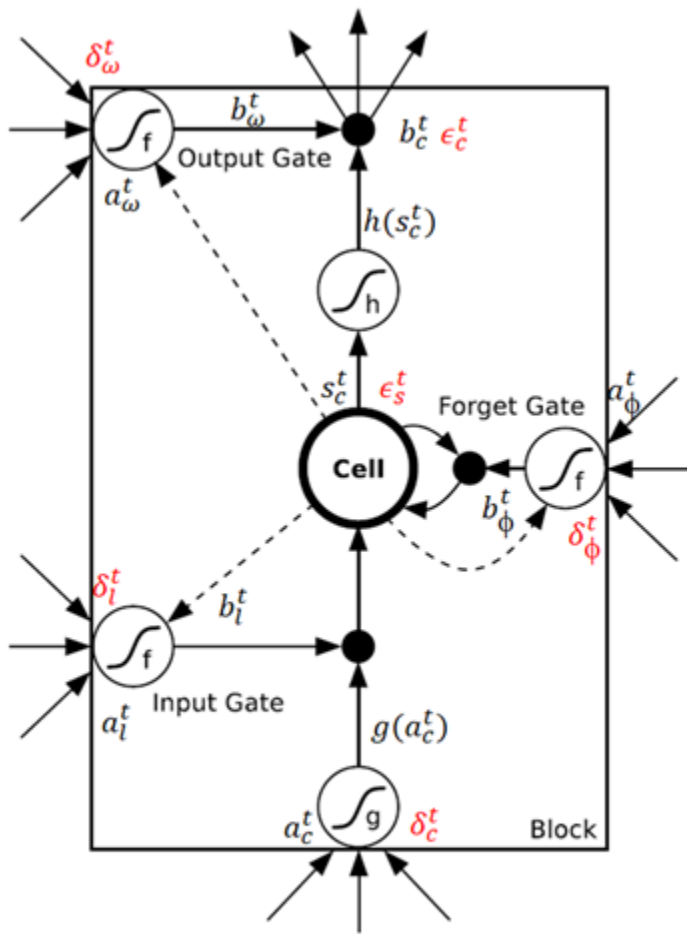


그림 6.12: 전방향 및 역방향 변수

Cells

$$\delta_c^t \equiv \frac{\partial L}{\partial a_c^t} = b_i^t g'(a_c^t) \epsilon_s^t \quad (6.4.4)$$

Forget Gates

$$\delta_\phi^t \equiv \frac{\partial L}{\partial a_\phi^t} = f'(a_\phi^t) \sum_{c=1}^C s_c^{t-1} \epsilon_s^t \quad (6.4.5)$$

Input Gates

$$\delta_i^t \equiv \frac{\partial L}{\partial a_i^t} = f'(a_i^t) \sum_{c=1}^C g(a_c^t) \epsilon_s^t \quad (6.4.6)$$



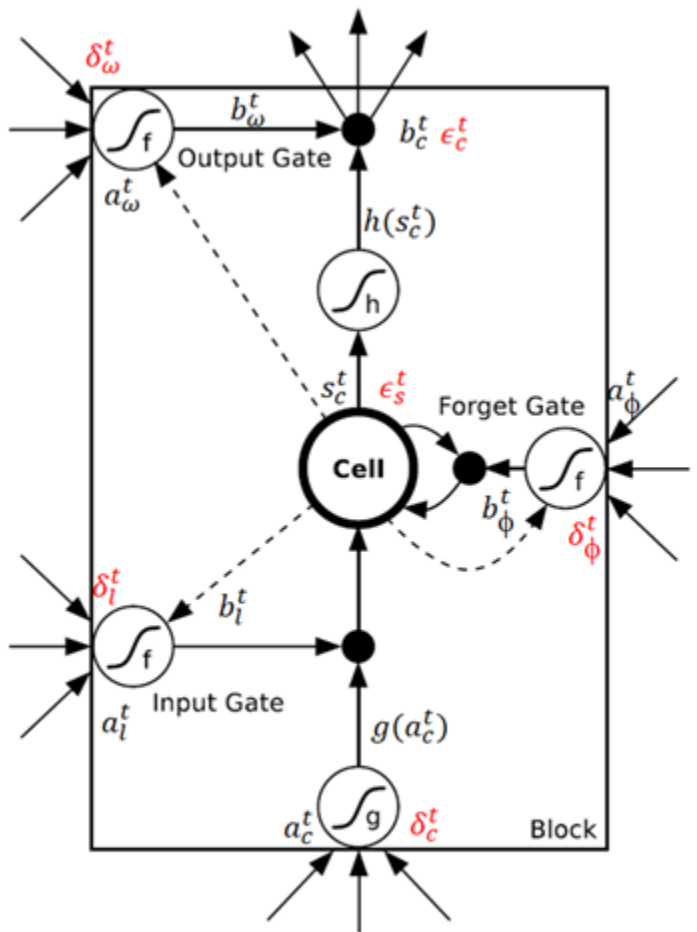


그림 6.12: 전방향 및 역방향 변수

$$\epsilon_c^t \stackrel{\text{def}}{=} \frac{\partial \mathcal{L}}{\partial b_c^t} \quad \epsilon_s^t \stackrel{\text{def}}{=} \frac{\partial \mathcal{L}}{\partial s_c^t}$$

Cell Outputs

$$\epsilon_c^t = \sum_{k=1}^K w_{ck} \delta_k^t + \sum_{g=1}^G w_{cg} \delta_g^{t+1}$$

Output Gates

$$\delta_\omega^t = f'(a_\omega^t) \sum_{c=1}^C h(s_c^t) \epsilon_c^t$$

States

$$\epsilon_s^t = b_\omega^t h'(s_c^t) \epsilon_c^t + b_\phi^{t+1} \epsilon_s^{t+1} + w_{ci} \delta_i^{t+1} + w_{c\phi} \delta_\phi^{t+1} + w_{cw} \delta_\omega^t$$

Cells

$$\delta_c^t = b_i^t g'(a_c^t) \epsilon_s^t$$

Forget Gates

$$\delta_\phi^t = f'(a_\phi^t) \sum_{c=1}^C s_c^{t-1} \epsilon_s^t$$

Input Gates

$$\delta_i^t = f'(a_i^t) \sum_{c=1}^C g(a_c^t) \epsilon_s^t$$