B Notation

Basic Concepts

Scalars: small *italic* letters.....*a,b,c*

Vectors: small **bold** nonitalic letters.....**a,b,c**

Matrices: capital **BOLD** nonitalic letters.....**A,B,C**

Language

Vector means a column of numbers.

Row vector means a row of a matrix used as a vector (column).

General Vectors and Transformations (Chapters 5 and 6)

$$x = A(y)$$

Weight Matrices

Scalar Element

$w^k_{i,j}(t)$

i - row, j - column, k - layer, t - time or iteration

Matrix

$\mathbf{W}^k(t)$

Column Vector

$\mathbf{w}^k_j(t)$

Row Vector

${}_i\mathbf{w}^k(t)$

Bias Vector

Scalar Element

$b^k_i(t)$

Vector

$\mathbf{b}^k(t)$

Input Vector

Scalar Element

$p_i(t)$

As One of a Sequence of Input Vectors

$\mathbf{p}(t)$

As One of a Set of Input Vectors

$\mathbf{p}_q$

Net Input Vector

Scalar Element

$n_i^k(t)$ or $n_{i,q}^k$

Vector

$\mathbf{n}^k(t)$ or $\mathbf{n}_q^k$

Output Vector

Scalar Element

$a_i^k(t)$ or $a_{i,q}^k$

Vector

$\mathbf{a}^k(t)$ or $\mathbf{a}_q^k$

Transfer Function

Scalar Element

$a_i^k = f^k(n_i^k)$

Vector

$\mathbf{a}^k = \mathbf{f}^k(\mathbf{n}^k)$

Target Vector

Scalar Element

$t_i(t)$ or $t_{i,q}$

Vector

$\mathbf{t}(t)$ or $\mathbf{t}_q$

Set of Prototype Input/Target Vectors

$\{\mathbf{p}_1, \mathbf{t}_1\}, \{\mathbf{p}_2, \mathbf{t}_2\}, \ldots, \{\mathbf{p}_Q, \mathbf{t}_Q\}$

Error Vector

Scalar Element

$e_i(t) = t_i(t) - a_i(t)$ or $e_{i,q} = t_{i,q} - a_{i,q}$

Vector

$\mathbf{e}(t)$ or $\mathbf{e}_q$

Sizes and Dimensions

Number of Layers, Number of Neurons per Layer

M, S^k

Number of Input Vectors (and Targets), Dimension of Input Vector

Q, R

Parameter Vector (includes all weights and biases)

Vector

$\mathbf{x}$

At Iteration k

$\mathbf{x}(k)$ or $\mathbf{x}_k$

Norm

$\|\mathbf{x}\|$

Performance Index

$F(\mathbf{x})$

Gradient and Hessian

$\nabla F(\mathbf{x}_k) = \mathbf{g}_k$ and $\nabla^2 F(\mathbf{x}_k) = \mathbf{A}_k$

Parameter Vector Change

$$\Delta \mathbf{x}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$$

Eigenvalue and Eigenvector

$$\lambda_i \text{ and } \mathbf{z}_i$$

Approximate Performance Index (single time step)

$$\hat{F}(\mathbf{x})$$

Transfer Function Derivative

Scalar

$$\dot{f}(n) = \frac{d}{dn} f(n)$$

Matrix

$$\dot{\mathbf{F}}^m(\mathbf{n}^m) = \begin{bmatrix} \dot{f}^m(n_1^m) & 0 & \ldots & 0 \\ 0 & \dot{f}^m(n_2^m) & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \dot{f}^m(n_{S^m}^m) \end{bmatrix}$$

Jacobian Matrix

$$\mathbf{J}(\mathbf{x})$$

Approximate Hessian Matrix

$$\mathbf{H} = \mathbf{J}^T \mathbf{J}$$

Sensitivity Vector

Scalar Element

$$s_i^m \equiv \frac{\partial \hat{F}}{\partial n_i^m}$$

Vector

$$\mathbf{s}^m \equiv \frac{\partial \hat{F}}{\partial \mathbf{n}^m}$$

Marquardt Sensitivity Matrix

Scalar Element

$$\tilde{s}_{i,h}^{m} \equiv \frac{\partial v_h}{\partial n_{i,q}^{m}} = \frac{\partial e_{k,q}}{\partial n_{i,q}^{m}}$$

Partial Matrix (single input vector $\mathbf{p}_q$) and Full Matrix (all inputs)

$$\tilde{\mathbf{S}}_q^m \text{ and } \tilde{\mathbf{S}}^m = \left[\tilde{\mathbf{S}}_1^m \; \tilde{\mathbf{S}}_2^m \; \ldots \; \tilde{\mathbf{S}}_Q^m\right]$$

Dynamic Networks

Sensitivity

$$s_{k,i}^{u,m}(t) \equiv \frac{\partial^e a_k^u(t)}{\partial n_i^m(t)}$$

Weight Matrices

$\mathbf{IW}^{m,l}(d)$ - input weight between input l and layer m at delay d

$\mathbf{LW}^{m,l}(d)$ - layer weight between layer l and layer m at delay d

Index Sets

$DL_{m,l}$ - delays in the tapped delay line between Layer l and Layer m.

$DI_{m,l}$ - delays in the tapped delay line between Input l and Layer m.

I_m - indices of input vectors that connect to layer m.

L_m^f - indices of layers that directly connect *forward* to layer m.

L_m^b - indices of layers that are directly connected backwards to layer m (or to which layer m connects forward) and that contain no delays in the connection.

$$E_{LW}^U(x) = \{u \in U \ni \exists(\mathbf{LW}^{x,u}(d) \neq 0, d \neq 0)\}$$

$$E_S^X(u) = \{x \in X \ni \exists(\mathbf{S}^{u,x} \neq 0)\}$$

$$E_S(u) = \{x \ni \exists(\mathbf{S}^{u,x} \neq 0)\}$$

$$E_{LW}^X(u) = \{x \in X \ni \exists(\mathbf{LW}^{x,u}(d) \neq 0, d \neq 0)\}$$

$$E_S^U(x) = \{u \in U \ni \exists(\mathbf{S}^{u,x} \neq 0)\}$$

Definitions

Input Layer (X) - has an input weight, or contains any delays with any of its weight matrices

Output Layer (U) - its output will be compared to a target during training, or it is connected to an input layer through a matrix that has delays associated with it.

Parameters for Backpropagation and Variations

Learning Rate and Momentum

α and γ

Learning Rate Increase, Decrease and Percentage Change

η , ρ and ζ

Conjugate Gradient Direction Adjustment Parameter

β_k

Marquardt Parameters

μ and ϑ

Generalization

Regularization Parameters

α , β and $\rho = \frac{\alpha}{\beta}$

Effective Number of Parameters

γ

Selected Model

M

Sum Squared Error and Sum Squared Weights

E_D , E_W

Maximum Likelihood and Most Probable Weights

$\mathbf{x}^{ML}$, $\mathbf{x}^{MP}$

Feature Map Terms

Distance Between Neurons

d_{ij} - distance between neuron i and neuron j

Neighborhood

$$N_i(d) = \{j, d_{ij} \le d\}$$

Grossberg and ART Networks

On-Center and Off-Surround Connection Matrices

$$^{+}\mathbf{W}^{1} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \text{ and } {}^{-}\mathbf{W}^{1} = \begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 0 \end{bmatrix}$$

Excitatory and Inhibitory Biases

$^{+}\mathbf{b}$ and $^{-}\mathbf{b}$

Time Constant

ε

Relative Intensity

$$\bar{p}_i = \frac{p_i}{P} \text{ where } P = \sum_{i=1}^{S^1} p_j$$

Instar and Outstar Weight Matrices

$\mathbf{W}^{1:2}$ and $\mathbf{W}^{2:1}$

Orienting Subsystem Parameters

$$\alpha, \beta \text{ and } \rho = \frac{\alpha}{\beta} \text{ (vigilance)}$$

ART1 Learning Law Parameter

ζ

Lyapunov Stability

Lyapunov Function

$V(\mathbf{a})$

Zero Derivative Set, Largest Invariant Set and Closure

Z, L and L°

Bounded Lyapunov Function Set

$\Omega_\eta = \{\mathbf{a}:\mathrm{V}(\mathbf{a}) < \eta\}$

Hopfield Network Parameters

Circuit Parameters

$T_{i,j}$, C, R_i, I_i, ρ

Amplifier Gain

γ