# THE EXPANSIONS AND APPLICATIONS OF GRONWALL – BELLMAN'S LEMMA

**Nguyen Dinh Phu and Hoang Thanh Long**
**Department of Mathematics & Informatics, University of Natural Sciences - VNU-HCM**
*(Received 02 November 2001, Revised 02 January 2002)*

**ABSTRACT:** The Gronwall – Bellman's Lemma (GBL) plays very important role in the researching Quality Theory of Ordinary Differential Equations. Beside the GBL there are many expansions, that was gave by some the authors, example Brézis[3], Bihari [2],...

This report has investigated two problems : Expansions and Applications of this important lemma. The special new results are given by the lemmas 6, 9 and many theorems for applied case, example the theorems 4, 5, 6 about stablity of similar control systems.

## §1. GRONWALL – BELLMAN'S LEMMA AND IT'S EXPANSIONS

### I. THE GRONWALL – BELLMAN'S LEMMA.

***1. Lemma 1:***

*Suppose that* $u(t) \ge 0$ *be differentiable on* $J = [t_o, t_1] \subset R$, $\square = u(t_o)$ , *and some real numbers* $k, c \in R$. *If u(t) satisfies:*

$$u'(t) \le c.\, u(t) + k, \quad \forall\, t \in J \tag{1.1}$$

*then*

$$u(t) \le \delta \exp\,[c(t - t_o)] + \frac{k}{c}\,\{\exp\,[c\,(t - t_o)] - 1\} \tag{1.2}$$

(There is Ordiginal Gronwall lemma)

***2. Lemma 2:***

*Suppose* $u(t) \ge 0$, $a(t) \ge 0$ on J, $k \ne 0$. *If* u(t), a(t)*, satisfy:*

$$u(t) \le a(t) + k\left|\int_{t_o}^{t} u(s)\,ds\right| \tag{1.3}$$

*then*

$$u(t) \le a(t) + k\int_{t_o}^{t} a(s)\exp\,[k(t - s)]\,ds\ \Big| \tag{1.4}$$

*Corollary 1:* In the case, when a(t) = a – const, u(t) satisfies (1.3), we have

$$u(t) \le a.\,\exp\,[k\,(t - t_o)]$$

### II. THE EXPANSIONS.

***3. Lemma 3:***

*Suppose* $u(t) \ge 0$, $\forall\, t \in J$, $a \ge 0$, $k \ge 0$ *and* $c > 0$ *real numbers. If* u(t) *satisfies:*

$$u(t) \le a + \int_{t_o}^{t} [cu(s) + k]\,ds \tag{1.5}$$

*Proof :* If we set $v(t) = a(t) + \int_{t_o}^{t} k(s)u(s)\,ds$, then $u(t) \le v(t)$

$v(t_o) = a(t_o)$. Because $a(t) \in C^1(J)$, then

$$v'(t) = a'(t) + k(t)\,u(t) \le a'(t) + k(t)\,v(t)$$

$$v'(t).\exp\left(-\int_{t_o}^{t} k(s)\,ds\right) \le \left[a'(t) + k(t)v(t)\right]\exp\left(-\int_{t_o}^{t} k(s)\,ds\right)$$

$$\frac{d}{dt}\left\{v(t)\exp\left(-\int_{t_o}^{t} k(s)\,ds\right)\right\} \le a'(t)\exp\left(-\int_{t_o}^{t} k(s)\,ds\right) \qquad (*)$$

Integrating (*) we have

$$v(t) \le a(t_o)\exp\left(\int_{t_o}^{t} k(s)\,ds\right) + \int_{t_o}^{t} a'(s)\exp\left(\int_{t_o}^{s} k(r)\,dr\right)ds \qquad (\square)$$

***7. Lemma 7 (Lemma Brézis [3]):***

*Suppose that* $u(t) \ge 0$, $a(t) > 0$, $b(t) > 0$, $k(t) > 0$ *for all* $t \in J$ *and satisfy:*

$$u(t) \le a(t) + b(t)\int_{t_o}^{t} k(s)\,u(s)\,ds \qquad (1.12)$$

*then* $$u(t) \le a(t) + b(t)\int_{t_o}^{t} a(s)\;k(s)\exp\left(\int_{s}^{t} k(r)\,b(r)\,dr\right)ds \qquad (1.13)$$

***8. Lemma 8:***

*Suppose that* $u(t) \ge 0$, $a(t) > 0$, *and* $K(t, s)$ *(where:* $|K(t, s)| \le M$, $K(t, s) > 0$, *for:* $t_o < s < t < t_1$ *and* $K(t, s) = 0$, *for :* $t_o < t < s < t_1$*) satisfy*

$$u(t) \le a(t) + \int_{t_o}^{t} K(t, s)\,u(s)\,ds \qquad (1.14)$$

*then* $u(t) \le \varphi(t)$, *where* $\varphi(t)$ *is a solution of integral equation :*

$$\varphi(t) = a(t) + \int_{t_o}^{t} K(t, s)\,\varphi(s)\,ds$$

*Proof:* Put $Ku = \int_{t_o}^{t} K(t, s)\,\psi(s)\,ds$ with

$$K^o u = u(t)\ ,\ K^n u = \int_{t_o}^{t} K(t,s)\,K^{n-1}\,u(s)\,ds$$

The operator K satisfies following process:

$$K^2 u = \int_{t_o}^{t} K(t,s)\,K^1 u(s)\,ds = \int_{t_o}^{t} K(t,s)\left(\int_{t_o}^{s} K(s,r)\,u(r)\,dr\right)ds$$

*then* $$u(t) \le a.\exp\left[c(t-t_o)\right] + \frac{k}{c}\left\{\exp\left[c(t-t_o)\right]-1\right\} \qquad (1.6)$$

*Proof :* Let u(t) satisfy (1.6), put $v(t) = \int_{t_o}^{t}\left[cu(s)+k\right]ds$.

such that $v(t_o) = 0$, $v'(t) = cu(t) + k \le cv(t) + k$ implies

$$v(t) \le v(t_o)\exp\left[c(t-t_o)\right] + \frac{k}{c}\left\{\exp\left[c(t-t_o)\right]-1\right\}$$

Because $v(t_o) = 0$ then $v(t) \le \frac{k}{c}\left\{\exp\left[c(t-t_o)\right]-1\right\}$

By in other hand, we have

$$u(t) - a \le \int_{t_o}^{t}\left\{\left[u(s)-a\right]+k+ca\right\}ds$$

If $z(t) = u(t) - a$, then $z(t) \le \int_{t_o}^{t}\left[cz(s)+k+ca\right]ds$ and implies

$$z(t) \le \frac{k+ca}{c}\left\{\exp\left[c(t-t_o)\right]-1\right\}$$

Finally, we have

$$u(t) = z(t) - a \le a\exp\left[c(t-t_o)\right] + \frac{k}{c}\left\{\exp\left[c(t-t_o)\right]-1\right\} \qquad (\square)$$

***4. Lemma 4:***

*Suppose* $u(t) \ge 0$, $k(t) \ge 0$, $\forall t \in J$ *and satisfy:*

$$u(t) \le a + \int_{t_o}^{t} k(s)u(s)\,ds \qquad (1.7)$$

*then* $$u(t) \le a\,.\exp\left(\int_{t_o}^{t} k(s)\,ds\right) \qquad (1.8)$$

***5. Lemma 5:***

*Suppose* $u(t) \ge 0$, $a(t) > 0$, $k(t) > 0$, $\forall t \in J$ *and satisfy:*

$$u(t) \le a(t) + \int_{t_o}^{t} k(s)u(s)\,ds \qquad (1.9)$$

*then* $$u(t) \le a(t) + \int_{t_o}^{t} a(s)k(s)\exp\left(\int_{s}^{t} k(r)\,dr\right)ds \qquad (1.10)$$

***6. Lemma 6***

*Suppose* $u(t) \ge 0$, $k(t) > 0$, $a(t) \in C^1(J)$, $a(t) > 0\ \forall t \in J$ *and satisfy (1.9), then*

$$u(t) \le a(t_o)\exp\left(\int_{t_o}^{t} k(r)\,dr\right) + \int_{t_o}^{t} a'(s)\exp\left(\int_{t_o}^{s} k(r)\,dr\right)ds \qquad (1.11)$$

The last formula (**) likes as (1.9). Using the lemma 5 we have the result.

***10. Lemma 10 (Lemma Bihari [2]):***

*Suppose* $u(t) \geq 0$, $f(t) \geq 0$, $\forall\, t \in [t_0, +\infty)$ *and satisfy:*

$$u(t) \leq c + \int_{t_0}^{t} f(s)\,\phi(u(s))\,ds \qquad (1.17)$$

*Where* $c > 0$, $\phi(u)$ *defined in* $0 < u < \bar{u} < +\infty$

*and*
$$\psi(u) = \int_{c}^{u} \frac{dv}{\phi(v)}, \quad 0 < u < \bar{u} \qquad (1.18)$$

*If*
$$\int_{t_0}^{t} f(s)\,ds < \psi(\bar{u}_{-0}) \qquad (1.19)$$

*then*
$$u(t) \leq \psi^{-1}\left[\int_{t_0}^{t} f(s)\,ds\right] \qquad (1.20)$$

## §2. THE APPLICATIONS

There are many applications of GBL and it's expansions, but in this report we limit the area of this applications.

### I. PROOF OF THE BASIS THEOREMS.

The GBL plays as instrument for proof of the basis theorems.

We consider two systems:

$$x' = X(t, x) \qquad (2.1)$$

$$y' = X(t, y) + R(t, y) \qquad (2.2)$$

Where $x, y \in B$ – Banach space, on D:

$$D = \{(x,t) : \|x - x_0\| \leq r,\ t \in J\}$$

The functions X(t, x) and R(t, x) satisfy:

$$\|X(t,x) - X(t,y)\| \leq L\|x - y\| \qquad (2.3)$$

$$\|R(t,y)\| \leq \theta(t),\ \theta(t) \in L^1(J) \qquad (2.4)$$

Using the GBL and it's expansions (Lemmas 3 – 10) we can proof the following theorems.

***Theorem 1:*** *If* $M_0 = \sup_D \|X(t,x)\|$, $M = M_0 + r.L^{-1}$ *then exists unique solution of system (2.1) for every* $t : |t - t_0| \leq r.\,M^{-1}$.

***Theorem 2:*** *If* x(t) *is a solution of (2.1) and y(t) is a solution of (2.2),* $\delta = \|y(t_0) - x(t_0)\|$ *then we have an estimation:*

$$\|y(t) - x(t)\| \leq \delta \exp[L(t - t_0)] + \int_{t_0}^{t} \exp[L(t - t_0)]\theta(s)\,ds$$

$$= \int_{t_0}^{t}\left(\int_{r}^{t} K(t,s).K(s,r)\,ds\right)u(r)\,dr = \int_{t_0}^{t} K_2(t,r)u(r)\,dr$$

$$K^n u = \int_{t_0}^{t} K_n(t,r)u(r)\,dr$$

Where $K_1(t, r) = K(t, r)$ , $K_n(t, r) = \int_{t_0}^{t} K(t,s)K_{n-1}(s,r)\,ds$

Estimate $|K_n(t, r)|$ and $|K^n u|$:

$$|K_n(t, r)| \le M^n \frac{(t-r)^{n-1}}{(n-1)!}$$

$$|K^n u| \le M^n \frac{(t-r)^{n-1}}{(n-1)!} . \int_{t_0}^{t} |u(s)|\,ds$$

When $n \to \infty$ we have $|K^n u| \to 0$

If $P_n(K) = I + K^1 + K^2 + ... + K^n$ then

$$u(t) \le P_n(K)\, a(t) + K^{n+1} u$$

When $n \to \infty$, then $K^{n+1} u \to 0$ and $P_n(K)\, a(t) \to \varphi(t)$ so $u(t) \le \varphi(t)$

Where $\varphi(t) = a(t) + \int_{t_0}^{t} K(t,s)\varphi(s)\,ds$ (□)

***9. Lemma 9:***

*Suppose* $u(t) \ge 0$, $a(t), b(t), k(t) > 0$, $\forall\, t \in J$ *and satisfy:*

i) $c(t) = \dfrac{a(t)}{b(t)} > 0$ *and* $\exists\, c'(t)$

ii) $u(t) \le a(t) + b(t) \int_{t_0}^{t} k(s)u(s)\,ds$ (1.12)

*then* $u(t) \le c(t_o)\, b(t)\, exp\left(\int_{t_0}^{t} z(s)\,ds\right) + b(t) \int_{t_0}^{t} c'(s) \exp\left(\int_{s}^{t} z(r)\,dr\right) ds$ (1.16)

*Where* $z(t) = b(t) . k(t)$.

*Proof :* From (1.12) we estimate

$$\frac{u(t)}{b(t)} \le \frac{a(t)}{b(t)} + \int_{t_0}^{t} k(s)u(s)\,ds$$

Where $\dfrac{u(t)}{b(t)} = y(t)$, $k(t)b(t) = z(t)$ and $\dfrac{a(t)}{b(t)} = c(t)$

Implies $y(t) \le c(t) + \int_{t_0}^{t} z(s)y(s)\,ds$ (**)

$$\leq k(s) \| x(s) - y(s) \| + p(s).$$

Finally, we have $\psi(t) < \delta + \int_{t_0}^{t} p(s)\,ds + \int_{t_0}^{t} k(s)\,\psi(s)\,ds$

or $\psi'(t) < k(t)\ \psi(t) + p(t)$. If function $\varphi(t)$ is a solution of differential equation

$$\varphi'(t) = k(t)\,\varphi(t) + p(t)$$

then $\psi(t) \leq \varphi(t)$, for every $t \in J$. (□)

## II. STABILITY OF THE SIMILAR CONTROL SYSTEMS.

### *1. The Linear Systems.*

Usually we consider system

$$x' = Ax + R(x) \tag{2.8}$$

as the first approximative systems with necessary condition

$$\lim_{\|x\|\to 0} \frac{\|R(x)\|}{\|x\|} = 0 \tag{2.9}$$

But we also can see R(x) as control function for process x(t) of Linear system.

$$x' = Ax \tag{2.10}$$

Where $A \in \aleph^n$ – matrix and

$$\lim_{\|x\|\to 0} \frac{\|R(x)\|}{\|x\|} \neq 0 \tag{2.11}$$

Particularly, we consider the case, when $R(x) = Bx$ with $\lim_{\|x\|\to 0} \frac{\|Bx\|}{\|x\|} \neq 0$,

that is

$$x' = Ax + Bx \tag{2.12}$$

***Theorem 4:*** *Suppose* $A, B \in \aleph^n$ *– matries , which satisfy*

i) $\mathrm{Re}\,\lambda_j(A) > 0$ *for many* $j \leq n$ *(*$\lambda_j$ *- Egeinvalue)*

ii) $C = A + B$ *with* $\mathrm{Re}\,\lambda_j(C) < 0$ *for* $\forall\, j = \overline{1,n}$ *then system (2.12) is stable.*

### *2. The Nonlinear Control:*

Beside the system (2.12) we consider system with nonlinear control:

$$x' = Ax + \varphi(t, x) \tag{2.13}$$

Where

$$\lim_{\|x\|\to 0} \frac{\|\varphi(t,x)\|}{\|x\|} \neq 0 \tag{2.14}$$

***Theorem 5 :*** *If* $\mathrm{Re}\,\lambda_j(A) < 0$ *for* $\forall\, j = \overline{1,n}$ *and :*

i) $\|\varphi(t,x)\| < \gamma(t)\|x\|$ (2.15)

*Corollary 2:*

Suppose $\theta(t) = N$ – constant

We have an estimation.

$$\|y(t) - x(t)\| \le \delta \exp[L(-t_o)] + \frac{N}{L}\{\exp[L(t - t_o)] - 1\}$$

- In the case, when $N = 0$:

$$\|y(t) - x(t)\| \le \delta \exp[L(t - t_o)]$$

- In the case, when $N > 0$, $\delta = 0$:

$$\|y(t) - x(t)\| \le \frac{N}{L}\{\exp[L(t - t_o)] - 1\}.$$

Nowaway, beside the system (2.1) we consider once more

$$x' = Y(t, x) \qquad (2.5)$$

and suppose that:

$$\|X(t, x) - X(t, y)\| \le k(t)\|x - y\| \qquad (2.6)$$

$$\|X(t, x) - Y(t, x)\| \le p(t). \qquad (2.7)$$

the functions $k(t), p(t) \in L^1(J)$

***Theorem 3:*** *If $x(t)$ is a solution of (2.1) with $x(t_o) = x_o$ and $y(t)$ is a solution of (2.5) with $y(t_o) = y_o \in U_\delta(x_o)$ then $\psi(t) = \|x(t) - y(t)\|$ satisfies:*

$$\psi(t) \le \delta + \int_{t_o}^{t} p(s)\,ds + \int_{t_o}^{t} k(s)\psi(s)\,ds$$

*and $\psi(t) \le \varphi(t)$ where $\varphi(t)$ is a solution of differential equation:* $\varphi' = k(t)\varphi + p(t)$.

*Proof:* We can write

$$x(t) = x(t_o) + \int_{t_o}^{t} X(s, x(s))\,dx$$

$$y(t) = y(t_o) + \int_{t_o}^{t} Y(s, y(s))\,ds$$

$$\psi(t) = \|x(t) - y(t)\| \le \|x(t_o) - y(t_o)\| + \int_{t_o}^{t} \|X(s, x(s)) - Y(x, y(s))\|\,ds$$

Because $y(t_o) = y_o \in U_\delta(x_o)$, then $\|x(t_o) - y(t_o)\| < \delta$

In other hand, we have

$$\|X(s, x(s)) - Y(s, y(s))\| = \|X(s, x(s)) - X(s, y(s)) + X(s, y(s)) - Y(s, y(s))\|$$

$$\le \|X(s, x(s)) - X(s, y(s))\| + \|X(s, y(s)) - Y(s, y(s))\|.$$

# MỞ RỘNG VÀ ỨNG DỤNG BỔ ĐỀ GRONWALL – BELLMAN

**Nguyễn Đình Phư và Hoàng Thanh Long**
**Khoa Toán – Tin học, Đại học KHTN, Đại học Quốc Gia TP.HCM**
*(Bài nhận ngày 02 tháng 11 năm 2001, hoàn chỉnh sửa chữa ngày 02 tháng 1 năm 2002)*

**TÓM TẮT:** Bổ đề Gronwall – Bellman (GBL) đóng vai trò quan trọng trong nghiên cứu lý thuyết định tính phương trình vi phân. Bên cạnh bổ đề GBL có một số mở rộng của một vài tác giả, ví dụ Brézis [3], Bihari [2], ...
Bài này khảo cứu hai vấn đề : mở rộng và ứng dụng bổ đề quan trọng này. Những kết quả đặc biệt mới được trình bày trong các bổ đề 6, 9 và một số định lý cho ứng dụng, ví dụ định lý 4, 5, 6 về ổn định của các hệ tựa điều khiển

## REFERENCE

[1] Barbasin E. A., Introduction into Stability Theory (In Russian).
Nauka Pub., Moscow 1967 (224 p.)

[2] Bihari J., A generalization of a lemma of Bellman and it's applications to uniqueness problem of Differential Equations.
Acta Math. Acad. Scient. Hungary, 1956, pp 81 – 89

[3] Brézis H., Operateurs Maximaux Monotones et Semi – groupes.
North Holland, 1973.

[4] Demidovich B. P., Course for Mathematical Stability Theory (In Russian)
Nauka Pub., Moscow 1967. (472 p.)

[5] Nguyen Dinh Phu, The Stability Theory and Applications.
Education Pub., Hà Nội 1996 (264 p.)

[6] Nguyen Dinh Phu, Ordinary Differential Equations
VNU – HCMC Pub., HoChiMinh 2001 (340 p.).

[7] Nguyen Dinh Phu, On the Thesis of stable condition of the Similar Control Systems.
Proceedings of the Conference on Natural Sciences HCM University, HCM City 1995 pp.78 – 81.

ii) $\int_0^\infty \gamma(d)\,dt < \infty$ (2.16)

*then system (2.13) is stable.*

*Proof :* We have a Cauchy operator for system (2.10) : $W(t, s) = X(t).\ X^{-1}(s)$ and a solution of system (2.13) is writed in Cauchy form :

$$x(t) = W(t, t_0)x_0 + \int_0^t W(t,s)\,\varphi(s, x(s))\,ds$$

that implies:

$$\|x(t)\| \le |W(t, t_0)|\,\|x_0\| + \int_0^t |W(t,s)|\,\|\varphi(s, x(s))\|\,ds$$

Because $\mathrm{Re}\,\lambda j\,(A) < 0,\ \forall\, j = \overline{1,n}$, then exist numbers $B$, $\alpha > 0$ such that

$$|W(t, t_0)| < B\ \exp[-\alpha\,(t - t_0)]$$

$$\|x(t)\| \le B\exp[-\alpha(t - t_0)] + \int_0^t B\exp[-\alpha(t - s)]\gamma(s)\,\|x(s)\|\,ds$$

Setting $\varphi(t) = \|x(t)\|\exp(\alpha t) > 0$, we have

$$\varphi(t) \le B\|x_0\|\exp(\alpha\, t_0) + B\int_{t_0}^t \gamma(s)\,\varphi(s)\,ds$$

Using lemma 4 of GBL'S expansions we get

$$\varphi(t) \le Be^{\alpha t_0}\|x_0\|\exp\left(\int_0^t \gamma(s)\,ds\right)$$

or $\|x(t)\| \le K.\exp(-\alpha t)$

that means systems (2.13) is stable (□)

***Theorem 6:*** *If the control function $\varphi(t, x)$ satisfies:*

$$\frac{\|\varphi(t,x)\|}{\|x\|} < \frac{\lambda_0}{2} \qquad (2.17)$$

*Where* $\lambda_0 = \min\limits_{i,j}|\lambda_i(A) + \lambda_j(A)|$ , *for* $i, j = \overline{1,n}$

*then system (2.13) is stable.*

*Corollary 3:* Suppose that A is stable matrix. If

i) or $\|\varphi(t,x)\| \le \|x\|^\beta\,\gamma(t)$, where $\beta > 0$, $\|x\|^{\beta-1} < \frac{\lambda_0}{2}$

ii) or $\|\varphi(t,x)\| \le h(\|x\|)\,\gamma(t)$, where $h(u) < \frac{\lambda_0}{2}$

then system (2.13) is stable.

$$\delta_{j,m} = \left( \sum_{k=1}^{n_{m+1}} W_{kj,m+1} \delta_{k,m+1} \right) . f'(N_{j,m})$$

So the weight changes are:

$$\Delta W_{ji,m}(n+1) = \eta \cdot \delta_{j,m} \cdot O_{i,m-1} \quad (11)$$

where

$$\delta_{j,m} = \begin{cases} f'(N_{j,m}).(t_j - O_{j,m}) & (12a) \\ f'(N_{j,m}). \sum_{k=1}^{n_{m+1}} W_{kj,m+1} \delta_{k,m+1} & (12b) \end{cases}$$

In Equation 12a, m is for output layer, meanwhile in Equation 12b, m is for hidden layer. In Equation 11, $\eta$ is a constant which represents learning rate. The larger $\eta$, the larger the changes in weight, thus the faster desired weight found. But if $\eta$ is too big, it causes an oscillation. The problem is to choose the maximum $\eta$ without leading to oscillation. To do this, RUMELHART et al. (1986) proposed an additional term called momentum that they believed would increase learning rate without leading to oscillation. With the addition of momentum term, weights are modified according to the following equations:

$$W_{ji,m}(n+1) = W_{ji,m}(n) + \Delta W_{ji,m}(n+1) \quad (13)$$

where $\Delta W_{ji,m}(n+1) = \eta \cdot \delta_{j,m} \cdot O_{i,m-1} + \alpha \cdot \Delta W_{ji,m}(n)$ (14)

with $\alpha$ being a constant which determines the effect of the past weight changes on the current direction of the movement. In practice, $\alpha$ is set around 0.9. The reference to DUC (2000) may illustrate the standard back propagation training algorithm.

## 4. Artificial Neural Network and Runoff Model

There are various models that can be used to simulate the monthly runoff process, namely HARMONIC, THOMAS and FIERING, and TANK. The first two models do not create a linkage between runoff and causal rainfall. Consequently, the accuracy of simulated results is limited due to the randomness of rainfall. The last model requires a large amount of physical factors. As a result, the complexity and costliness in data collection and computation may be encountered. Contrarily, the artificial neural network is a fuzzy model to link runoff and its causal rainfall. It requires more data sources than the above-mentioned models do. The data include monthly runoffs at considered location and upstream stations, monthly rainfalls and evaporation, etc. at surrounding stations within the basin. The monthly runoff data at upstream station has normally a very good correlation with that at downstream station. As such, the randomness of rainfall process that effects to the runoff generation (deterministic case) may be avoided. This makes the artificial neural network model may be better for runoff simulation if the suitable selection of inputs and targets is obtained.

## 5. Application of the Artificial Neural Network Model to Simulate Monthly Runoff

### *5.1 Standards for statistic performance:*

To evaluate the simulated results, the standards for statistic performance including Efficiency Index (EI), Root Mean Square Error (RMSE), Root Mean Square Error Mean (RMSEM), Mean Absolute Deviation (MAD), Root Mean Square Error over Standard Deviation (RMSES) were used in this study. The reference to DUC (2000) may give detailed explanation for these standards.

$$O_{j,m} = f(N_{j,m}) = \frac{1-e^{-N_{j,m}}}{1+e^{-N_{j,m}}} \quad (4a)$$

with $f'(N_{j,m}) = 0.5\,(1-O^2_{j,m})$ (5a)

can also be used as semi-linear function.

Now, the output $O_{j,m}$ is used as the input at layer m for the next layer m+1, that means

$I_{j,m} = O_{j,m}$ (3a)

The forward process is repeated until the output signal at output target layer is reached.

If the fitting between final outputs and the targets at respective units of target layer attains then the process stops. Otherwise, the search of ANN structure has to be redone until errors between the final outputs and targets are small as appropriate for all the patterns of respective input and target values.

$$E_p = \frac{1}{2}\sum_{j=1}^{n_L}(t_j - O_{j,L})^2 \quad (6)$$

Whereas total error is:

$$E = \sum_p E_p \quad (7a)$$

$$E \le \varepsilon \quad (7b)$$

where, ε is the allowable total sum square error.

For this training process, several algorithms can be done. In this study the back propagation neural network algorithm is used.

### 3. Back Propagation Neural Network Algorithm

Back propagation method tries to minimize the total error (E) by adjusting its weights:

$$\frac{\partial E_p}{\partial W_{ji,m}} = \frac{\partial E_p}{\partial N_{j,m}} \cdot \frac{\partial N_{j,m}}{\partial W_{ji,m}} \quad (8)$$

$$\frac{\partial E_p}{\partial W_{ji,m}} = -\delta_{j,m} . O_{i,m-1}$$

where $\delta_{j,m} = -\frac{\partial E_p}{\partial N_{j,m}}$

For units in the output layer:

$$\delta_{j,m} = -\frac{\partial E_p}{\partial O_{j,m}} \cdot \frac{\partial O_{j,m}}{\partial N_{j,m}} \quad (9)$$

$\delta_{j,m} = (t_j - O_{j,m}) \,.\, f'(N_{j,m})$

For units in the hidden layers:

$$\delta_{j,m} = -\frac{\partial E_p}{\partial O_{j,m}} \cdot \frac{\partial O_{j,m}}{\partial N_{j,m}}$$

$$\delta_{j,m} = \left( -\sum_{k=1}^{n_{m+1}} \frac{\partial E_p}{\partial N_{k,m+1}} \frac{\partial N_{k,m+1}}{\partial O_{j,m}} \right) \cdot \frac{\partial O_{j,m}}{\partial N_{j,m}} \quad (10)$$

combination of input and output nodes and inside structure will reduce significantly time consumption during training process. Figure 1 shows this simple ANN structure. Connection from node (i) in layer m-1, to another node (j), in consecutive layer m, is a weight ($W_{ji,m}$). $I_i$ is called as the signal at node i of input layer; $O_{j,m}$ as output signal of unit j at layer m (may be hidden or output layer); $n_0$ , $n_1$ and $n_2$ are the numbers of unit in input, hidden and output layers, respectively. $t_j$ is called as the target of unit j at output layer. Then the coming signal (called as NET) at node j can be calculated by the following summation:

$$N_{j,m} = \sum_{i=0}^{n_{m-1}} W_{ji,m} \cdot O_{i,m-1} \quad ; \quad m = 1, 2 \tag{1}$$

where, if $i \neq 0$ and m=1 then $O_{i,0} = I_i$ : input unit (2)

if $i = 0$ then $O_{0,m-1} = 1$ and $W_{j0,m} = \theta_{j,m}$ : a bias

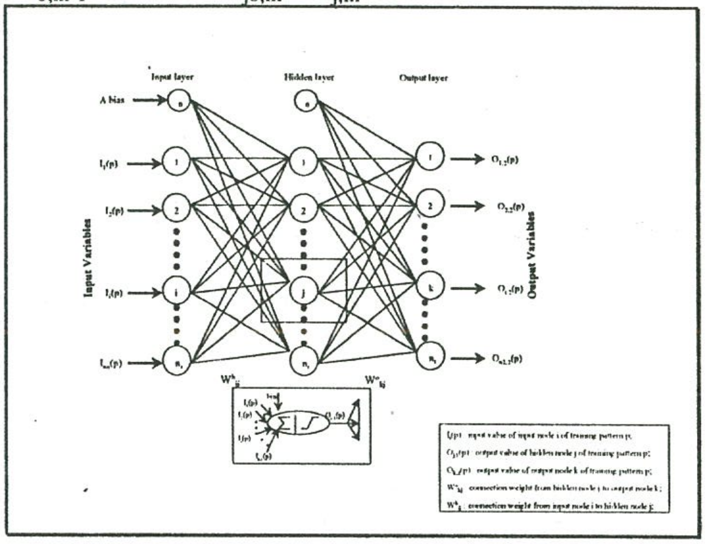


Figure 1 Typical Three-Layer Feed Forward Artificial Neural Network

After NET is calculated, an activation function f is applied to modify it, thereby producing the signal OUT (WASSERMAN, 1989) as follows:

$$O_{j,m} = f(N_{j,m}) \tag{3}$$

Function f was proposed as a semi-linear function (non-decreasing and differentiable to the net total output). In practice, a logistic activation function in [0,1]

$$O_{j,m} = f(N_{j,m}) = \frac{1}{1+e^{-N_{j,m}}} \tag{4}$$

is used as semi-linear function because it is non-decreasing function and has simple derivative:

$$f'(N_{j,m}) = O_{j,m}(1-O_{j,m}) \tag{5}$$

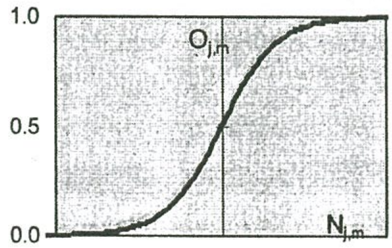


Figure 2 Activation Function in [0,1]

Another type of activation function in [-1,+1]