Expectation and Variance Derivation*

Mark Stefanski, Kanishka Bhaduri, Ashok N. Srivastava

March 9, 2010

The following two equations transform the input into output in a neural net:

$$\begin{aligned} O(x) &= B + QH(x) \\ H(x) &= \tanh(A + Wx) \end{aligned}$$

where x is an n-tuple of data represented as a column vector and the capital letters denote matrices. The dimensions of these matrices are:

$$\begin{array}{cc} W & m \times n \\ A & m \times 1 \\ Q & p \times m \\ B & p \times 1 \end{array}$$

and it is assumed that the entries of the matrices are taken from a normal distribution with mean 0. For simplicity, we assume $A + Wx$ is sufficiently close to 0 to permit the simplifying assumption that tanh is linear with constant slope a. Thus:

$$\begin{aligned} O(x) &= B + QH(x) \\ H(x) &= a(A + Wx) \end{aligned}$$

Combining these two formulae into a more convenient form:

$$O(x) = B + aQA + aQWx$$

*Supplemental material for the paper submitted to IEEE SMC part B.

Now we write

$$\begin{aligned} y_1 &= O(x_1) = B + aQA + aQWx_1 \\ y_2 &= O(x_2) = B + aQA + aQWx_2 \end{aligned}$$

Hence,

$$\Delta y = y_2 - y_1 = aQW(x_2 - x_1)$$

We proceed to derive $E[||y_2 - y_1||^2]$ and $Var[||y_2 - y_1||^2]$. To this end, we introduce $R = QW$ whose entries $r_{i.j}$ are given by

$$r_{i,j} = q_{i,1}w_{1,j} + q_{i,2}w_{2,j} + ... + q_{i,m}w_{m,j} = \sum_{t=1}^{m} q_{i,t}w_{t.j}$$

In turn,

$$\Delta y = aQW(x_2 - x_1)$$

whose entries are given by

$$\Delta y_j = a\sum_{s=1}^{n} r_{i,s}\Delta x_s = a\sum_{s=1}^{n}\sum_{t=1}^{m} q_{i,t}w_{t,s}\Delta x_s$$

It suffices to find $E[y_j^2]$ for $1 \leq j \leq p$ because

$$E[||\Delta y||^2] = E[\Delta y_1^2 + \Delta y_2^2 + ... + \Delta y_p^2] = E[\Delta y_1^2] + E[\Delta y_2^2] + ... + E[\Delta y_p^2]$$

So we write

$$\Delta y_j^2 = (a\sum_{s=1}^{n}\sum_{t=1}^{m} q_{i,t}w_{t,s}\Delta x_s)^2$$

In expanding Δy_j^2 we may ignore terms that are linear in $q_{i,t}$ or $w_{t,j}$, as these are independent and the expected value of each evaluates to 0. That is, the terms of interest have the form

$q_{i,t}^2 w_{t,j}^2$. Thus,

$$\begin{aligned}
E[\Delta y_j^2] &= a^2 E[(q_{i,q}w_{1,1} + ... + q_{i,m}w_{m,1})^2 \Delta x_1^2 + ... + (q_{i,1}w_{1,n} + ... + q_{i,m}w_{m,n})^2 \Delta x_n^2] \\
&= a^2 E[(q_{i,1}^2 w_{1,1}^2 + ... + q_{i,m}^2 w_{m,1}^2)\Delta x_1^2 + ... + (q_{i,m}^2 w_{1,n}^2 + ... + q_{i,m}^2 w_{m,n}^2)\Delta x_n^2] \\
&= a^2 (E[q_{i,1}^2 w_{1,1}^2 + ... + q_{i,m}^2 w_{m,1}^2]\Delta x_1^2 + ... + E[q_{i,m}^2 w_{1,n}^2 + ... + q_{i,m}^2 w_{m,n}^2]\Delta x_n^2) \\
&= a^2((E[q_{i,1}^2 w_{1,1}^2] + ... + E[q_{i,m}^2 w_{m,1}^2])\Delta x_1^2 \\
&\quad +... + (E[q_{i,m}^2 w_{1,n}^2] + ... + E[q_{i,m}^2 w_{m,n}^2])\Delta x_n^2) \\
&= a^2((E[q_{i,1}^2]E[w_{1,1}^2] + ... + E[q_{i,m}^2]E[w_{m,1}^2])\Delta x_1^2 \\
&\quad +... + (E[q_{i,m}^2]E[w_{1,n}^2] + ... + E[q_{i,m}^2]E[w_{m,n}^2])\Delta x_n^2) \\
&= a^2((\sigma_q^2\sigma_w^2 + ... + \sigma_q^2\sigma_w^2)\Delta x_1^2 + ... + (\sigma_q^2\sigma_w^2 + ... + \sigma_q^2\sigma_w^2)\Delta x_n^2) \\
&= a^2(m\sigma_q^2\sigma_w^2\Delta x_1^2 + ... + m\sigma_q^2\sigma_w^2\Delta x_n^2) \\
&= a^2 m\sigma_q^2\sigma_w^2 \sum_{j=1}^{n} \Delta x_j^2 \\
&= ma^2\sigma_q^2\sigma_w^2 ||\Delta x||^2
\end{aligned}$$

Recalling that

$$E[||\Delta y||^2] = E[\Delta y_1^2] + E[\Delta y_2^2] + ... + E[\Delta y_p^2]$$

We have

$$E[||\Delta y||^2] = pma^2\sigma_q^2\sigma_w^2 ||\Delta x||^2$$

It easily follows that

$$\begin{aligned}
E[||\Delta y||^2 - ||\Delta x||^2] &= E[||\Delta y||^2] - E[||\Delta x||^2] \\
&= pma^2\sigma_q^2\sigma_w^2 ||\Delta x||^2 - ||\Delta x||^2 \\
&= ||\Delta x||^2 (pma^2\sigma_q^2\sigma_w^2 - 1)
\end{aligned}$$

Now to compute $Var[||\Delta y||^2]$. We begin by noting that

$$\begin{aligned}
Var[||\Delta y||^2] &= Var[\Delta y_1^2 + \Delta y_2^2 + ...\Delta y_p^2] \\
&= \sum_{i=1}^{p}\sum_{j=1}^{p} Cov(\Delta y_i^2, \Delta y_j^2) \\
&= \sum_{i=1}^{p}\sum_{j=1}^{p} (E[\Delta y_i^2 \Delta y_j^2] - E[\Delta y_i^2]E[\Delta y_j^2]) \\
&= \sum_{i=1}^{p}\sum_{j=1}^{p} (E[\Delta y_i^2 \Delta y_j^2] - ma^2\sigma_q^2\sigma_w^2||\Delta x||^2)^2) \\
&= \sum_{i=1}^{p}\sum_{j=1}^{p} (E[\Delta y_i^2 \Delta y_j^2] - (m^2a^4\sigma_q^4\sigma_w^4||\Delta x||^4)
\end{aligned}$$

So it suffices to find $E[\Delta y_i^2 \Delta y_j^2]$, which we do separately for $i = j$ and $i \neq j$. Recalling that

$$r_{i.j} = q_{i,1}w_{1,j} + q_{i,2}w_{2,j} + ... + q_{i,m}w_{m,j} = \sum_{t=1}^{m} q_{i,t}w_{t.j}$$

we note that

$$\begin{aligned}
\Delta y_i^2 &= (a\sum_{s=1}^{n}\sum_{t=1}^{m} q_{i,t}w_{t,s}\Delta x_s^2)^2 \\
&= (a\sum_{s=1}^{n} r_{i,s}\Delta x_s^2)^2 \\
&= a^2(\sum_{s=1}^{n} r_{i,s}\Delta x_s^2)^2 \\
&= a^2((r_{i,1}^2\Delta x_1^2 + r_{i,1}r_{i,2}\Delta x_1\Delta x_2 + ... + r_{i,1}r_{i,n}\Delta x_1\Delta x_n) \\
&\quad +... + (r_{i,n}r_{i,1}\Delta x_n\Delta x_1 + r_{i,n}r_{i,2}\Delta x_n\Delta x_2 + ... + r_{i,n}^2\Delta x_n^2))
\end{aligned}$$

Similarly,

$$\begin{aligned}
y_j^2 &= a^2((r_{j,1}^2\Delta x_1^2 + r_{j,1}r_{j,2}\Delta x_1\Delta x_2 + ... + r_{j,1}r_{j,n}\Delta x_1\Delta x_n) \\
&\quad +... + (r_{j,n}r_{j,1}\Delta x_n\Delta x_1 + r_{j,n}r_{j,2}\Delta x_n\Delta x_2 + ... + r_{j,n}^2\Delta x_n^2))
\end{aligned}$$

So the terms of $\Delta y_i^2\Delta y_j^2$ have the form

$$a^4(r_{i,a}r_{i,b}r_{j,c}r_{j,d})(\Delta x_a\Delta x_b\Delta x_c\Delta x_d)$$

where a, b, c, d are indices. In turn, the terms of $r_{i,a}r_{i,b}r_{j,c}r_{j,d}$ have the form

$$(q_{i,s}w_{s,a})(q_{i,t}wt, b)(q_{j,u}w_{u,c})(q_{j,v}wv, d) = (q_{i,s}q_{i,t}q_{j,u}q_{j,v})(w_{s,a}w_{t,b}w_{u,c}w_{v,d})$$

So the terms of $\Delta y_i^2 \Delta y_j^2$ have the form

$$a^4(q_{i,s}q_{i,t}q_{j,u}q_{j,v})(w_{s,a}w_{t,b}w_{u,c}w_{v,d})(\Delta x_a \Delta x_b \Delta x_c \Delta x_d)$$

Yet again, we need only consider terms that are not linear in q or w because all others evaluate to zero under the expected value operator. So we consider all of the cases in which the terms do not evaluate to zero. First,

Table 1: default

i $\neq$ j	s = t $\neq$ u = v	s = t, u = v	a = b, c = d
	s = t $\neq$ u = v	s = u, t = v	a = c, b = d
	s = t $\neq$ u = v	s = v, t = u	a = d, b = c
	s = t = u = v		a = b $\neq$ c = d
	s = t = u = v		a = c $\neq$ b = d
	s = t = u = v		a = d $\neq$ b = c
	s = t = u = v		a = b = c = d

Second,

This enumeration of combinations reduces to

and

Table 2: default

$i = j$	$s = t \neq u = v$	$s = t,\ u = v$	$a = b,\ c = d$
	$s = t \neq u = v$	$s = u,\ t = v$	$a = c,\ b = d$
	$s = t \neq u = v$	$s = v,\ t = u$	$a = d,\ b = c$
	$s = u \neq t = v$	$s = t,\ u = v$	$a = b,\ c = d$
	$s = u \neq t = v$	$s = u,\ t = v$	$a = c,\ b = d$
	$s = u \neq t = v$	$s = v,\ t = u$	$a = d,\ b = c$
	$s = v \neq t = u$	$s = t,\ u = v$	$a = b,\ c = d$
	$s = v \neq t = u$	$s = u,\ t = v$	$a = c,\ b = d$
	$s = v \neq t = u$	$s = v,\ t = u$	$a = d,\ b = c$
	$s = t = u = v$		$a = b \neq c = d$
	$s = t = u = v$		$a = c \neq b = d$
	$s = t = u = v$		$a = d \neq b = c$
	$s = t = u = v$		$a = b = c = d$

Table 3: default

1	$i \neq j$	$s = t \neq u = v$	$a = b \neq c = d$
2		$s = t \neq u = v$	$a = b = c = d$
3		$s = t = u = v$	$a = b \neq c = d$
4		$s = t = u = v$	$a = c \neq b = d$
5		$s = t = u = v$	$a = d \neq b = c$
6		$s = t = u = v$	$a = b = c = d$

Table 4: default

1	i = j	s = t ≠ u = v	a = b ≠ c = d
2		s = t ≠ u = v	a = b = c = d
3		s = u ≠ t = v	a = c ≠ b = d
4		s = u ≠ t = v	a = b = c = d
5		s = v ≠ t = u	a = d ≠ b = c
6		s = v ≠ t = u	a = b = c = d
7		s = t = u = v	a = b ≠ c = d
8		s = t = u = v	a = c ≠ b = d
9		s = t = u = v	a = d ≠ b = c
10		s = t = u = v	a = b = c = d

Now for 1:

$$
\begin{aligned}
E[Type1] &= E[\sum_{s=1}^{m}\sum_{u=1,u\neq s}^{m}\sum_{a=1}^{n}\sum_{c=1,c\neq a}^{n} a^4 q_{i,s}^2 q_{j,u}^2 w_{s,a}^2 w_{u,c}^2 \Delta x_a^2 \Delta x_c^2] \\
&= a^4 \sum_{s=1}^{m}\sum_{u=1,u\neq s}^{m}\sum_{a=1}^{n}\sum_{c=1,c\neq a}^{n} E[q_{i,s}^2]E[q_{j,u}^2]E[w_{s,a}^2]E[w_{u,c}^2]\Delta x_a^2 \Delta x_c^2 \\
&= a^4 \sum_{s=1}^{m}\sum_{u=1,u\neq s}^{m}\sum_{a=1}^{n}\sum_{c=1,c\neq a}^{n} \sigma_q^2\sigma_q^2\sigma_w^2\sigma_w^2 \Delta x_a^2 \Delta x_c^2 \\
&= a^4\sigma_q^4\sigma_w^4 \sum_{s=1}^{m}\sum_{u=1,u\neq s}^{m}\sum_{a=1}^{n} \Delta x_a^2 \sum_{c=1,c\neq a}^{n} \Delta x_c^2 \\
&= a^4\sigma_q^4\sigma_w^4 \sum_{s=1}^{m}\sum_{u=1,u\neq s}^{m} \left(\sum_{a=1}^{n}\Delta x_a^2 \sum_{c=1}^{n}\Delta x_c^2 - \sum_{i=1}^{n}\Delta x_i^4\right) \\
&= a^4\sigma_q^4\sigma_w^4 \sum_{s=1}^{m}\sum_{u=1,u\neq s}^{m} \left(||\Delta x||^2||\Delta x||^2 - \sum_{i=1}^{n}\Delta x_i^4\right) \\
&= m(m-1)a^4\sigma_q^4\sigma_w^4 \left(||\Delta x||^4 - \sum_{i=1}^{n}\Delta x_i^4\right)
\end{aligned}
$$

For Type 2:

$$
\begin{aligned}
E[Type2] &= E[\sum_{s=1}^{m}\sum_{u=1,u\neq s}^{m}\sum_{a=1}^{n} a^4 q_{i,s}^2 q_{j,u}^2 w_{s,a}^4 \Delta x_a^4] \\
&= a^4 \sum_{s=1}^{m}\sum_{u=1,u\neq s}^{m}\sum_{a=1}^{n} E[q_{i,s}^2]E[q_{j,u}^2]E[w_{s,a}^4]\Delta x_a^4 \\
&= a^4 \sum_{s=1}^{m}\sum_{u=1,u\neq s}^{m}\sum_{a=1}^{n} \sigma_q^2\sigma_q^2(3\sigma_w^4)\Delta x_a^4 \\
&= 3a^4\sigma_q^4\sigma_w^4 \sum_{s=1}^{m}\sum_{u=1,u\neq s}^{m}\sum_{a=1}^{n} \Delta x_a^4 \\
&= 3m(m-1)a^4\sigma_q^4\sigma_w^4 \left(\sum_{a=1}^{n}\Delta x_a^4\right)
\end{aligned}
$$

For Type 3:

$$
\begin{aligned}
E[Type1] &= E[\sum_{s=1}^{m}\sum_{a=1}^{n}\sum_{c=1,c\neq a}^{n} a^4 q_{i,s}^2 q_{j,s}^2 w_{s,a}^2 w_{s,c}^2 \Delta x_a^2 \Delta x_c^2] \\
&= a^4 \sum_{s=1}^{m}\sum_{a=1}^{n}\sum_{c=1,c\neq a}^{n} E[q_{i,s}^2]E[q_{j,s}^2]E[w_{s,a}^2]E[w_{s,c}^2]\Delta x_a^2\Delta x_c^2 \\
&= a^4 \sum_{s=1}^{m}\sum_{a=1}^{n}\sum_{c=1,c\neq a}^{n} \sigma_q^2\sigma_q^2\sigma_w^2\sigma_w^2\Delta x_a^2\Delta x_c^2 \\
&= a^4\sigma_q^4\sigma_w^4 \sum_{s=1}^{m}\sum_{a=1}^{n}\Delta x_a^2 \sum_{c=1,c\neq a}^{n}\Delta x_c^2 \\
&= a^4\sigma_q^4\sigma_w^4 \sum_{s=1}^{m}\left(\sum_{a=1}^{n}\Delta x_a^2\sum_{c=1}^{n}\Delta x_c^2 - \sum_{i=1}^{n}\Delta x_i^4\right) \\
&= a^4\sigma_q^4\sigma_w^4 \sum_{s=1}^{m}\left(||\Delta x||^2||\Delta x||^2 - \sum_{i=1}^{n}\Delta x_i^4\right) \\
&= ma^4\sigma_q^4\sigma_w^4 \left(||\Delta x||^4 - \sum_{i=1}^{n}\Delta x_i^4\right)
\end{aligned}
$$

By symmetry, we have

$$E[Type3] = E[Type4] = E[Type5]$$

Finally, for Type 6:

$$\begin{aligned}
E[Type6] &= E[\sum_{s=1}^{m}\sum_{a=1}^{n} a^4 q_{i,s}^2 q_{j,s}^2 w_{s,a}^4 \Delta x_a^4] \\
&= a^4 \sum_{s=1}^{m}\sum_{a=1}^{n} E[q_{i,s}^2]E[q_{j,s}^2]E[w_{s,a}^4]\Delta x_a^4 \\
&= a^4 \sum_{s=1}^{m}\sum_{a=1}^{n} \sigma_q^2\sigma_q^2(3\sigma_w^4)\Delta x_a^4 \\
&= 3a^4\sigma_q^4\sigma_w^4 \sum_{s=1}^{m}\sum_{a=1}^{n} \Delta x_a^4 \\
&= 3ma^4\sigma_q^4\sigma_w^4 \sum_{a=1}^{n} \Delta x_a^4
\end{aligned}$$

So for $i = j$, $E[\Delta y_i^2 \Delta y_j^2]$ is given by

$$\begin{aligned}
E[\Delta y_i^4] &= E[Type1] + E[Type2] + 3E[Type3] + E[Type6] \\
&= m(m-1)a^4\sigma_q^4\sigma_w^4 \left(||\Delta x||^4 - \sum_{i=1}^{n} \Delta x_i^4 \right) \\
&\quad +3m(m-1)a^4\sigma_q^4\sigma_w^4 \left(\sum_{a=1}^{n} \Delta x_a^4 \right) \\
&\quad +3\left(ma^4\sigma_q^4\sigma_w^4 \left(||\Delta x||^4 - \sum_{i=1}^{n} \Delta x_i^4 \right) \right) \\
&\quad +3ma^4\sigma_q^4\sigma_w^4 \sum_{a=1}^{n} \Delta x_a^4 \\
&= m(m-1)a^4\sigma_q^4\sigma_w^4 \left(||\Delta x||^4 + 2\sum_{i=1}^{n} \Delta x_i^4 \right) + 3ma^4\sigma_q^4\sigma_w^4 ||\Delta x||^4
\end{aligned}$$

By symmetry, we have

$$E[Type1] = E[Type1] = E[Type3] = E[Type5]$$

By symmetry, we have

$$E[Type2] = E[Type2] = E[Type4] = E[Type6]$$

For Type 7:

$$
\begin{aligned}
E[Type7] &= E[\sum_{s=1}^{m}\sum_{a=1}^{n}\sum_{c=1,c\neq a}^{n} a^4 q_{i,s}^4 w_{s,a}^2 w_{s,c}^2 \Delta x_a^2 \Delta x_c^2] \\
&= a^4 \sum_{s=1}^{m}\sum_{a=1}^{n}\sum_{c=1,c\neq a}^{n} E[q_{i,s}^4]E[w_{s,a}^2]E[w_{s,c}^2]\Delta x_a^2 \Delta x_c^2 \\
&= a^4 \sum_{s=1}^{m}\sum_{a=1}^{n}\sum_{c=1,c\neq a}^{n} (3\sigma_q^2)\sigma_w^2\sigma_w^2 \Delta x_a^2 \Delta x_c^2 \\
&= 3a^4\sigma_q^4\sigma_w^4 \sum_{s=1}^{m}\sum_{a=1}^{n} \Delta x_a^2 \sum_{c=1,c\neq a}^{n} \Delta x_c^2 \\
&= 3a^4\sigma_q^4\sigma_w^4 \sum_{s=1}^{m}\left(\sum_{a=1}^{n}\Delta x_a^2 \sum_{c=1}^{n}\Delta x_c^2 - \sum_{i=1}^{n}\Delta x_i^4\right) \\
&= 3a^4\sigma_q^4\sigma_w^4 \sum_{s=1}^{m}\left(||\Delta x||^2||\Delta x||^2 - \sum_{i=1}^{n}\Delta x_i^4\right) \\
&= 3ma^4\sigma_q^4\sigma_w^4 \left(||\Delta x||^4 - \sum_{i=1}^{n}\Delta x_i^4\right)
\end{aligned}
$$

By symmetry, we have

$$E[Type7] = E[Type8] = E[Type9]$$

Finally, for Type 10:

$$
\begin{aligned}
E[Type10] &= E[\sum_{s=1}^{m}\sum_{a=1}^{n} a^4 q_{i,s}^4 w_{s,a}^4 \Delta x_a^4] \\
&= a^4 \sum_{s=1}^{m}\sum_{a=1}^{n} E[q_{i,s}^2]E[q_{j,s}^2]E[w_{s,a}^4]\Delta x_a^4 \\
&= a^4 \sum_{s=1}^{m}\sum_{a=1}^{n} (3\sigma_q^2)(3\sigma_w^4)\Delta x_a^4 \\
&= 9a^4\sigma_q^4\sigma_w^4 \sum_{s=1}^{m}\sum_{a=1}^{n} \Delta x_a^4 \\
&= 9ma^4\sigma_q^4\sigma_w^4 \sum_{a=1}^{n} \Delta x_a^4
\end{aligned}
$$

So for $i \neq j$, $E[\Delta y_i^2 \Delta y_j^2]$ is given by

$$\begin{aligned}
E[\Delta y_i^2 \Delta y_j^2] &= 3E[Type1] + 3E[Type2] + 3E[Type7] + E[Type10] \\
&= 3\left(m(m-1)a^4\sigma_q^4\sigma_w^4\left(||\Delta x||^4 - \sum_{i=1}^{n}\Delta x_i^4\right)\right) \\
&\quad +3\left(3m(m-1)a^4\sigma_q^4\sigma_w^4\left(\sum_{a=1}^{n}\Delta x_a^4\right)\right) \\
&\quad +3\left(3ma^4\sigma_q^4\sigma_w^4\left(||\Delta x||^4 - \sum_{i=1}^{n}\Delta x_i^4\right)\right) \\
&\quad +9ma^4\sigma_q^4\sigma_w^4\sum_{a=1}^{n}\Delta x_a^4 \\
&= 3m(m-1)a^4\sigma_q^4\sigma_w^4\left(||\Delta x||^4 + 2\sum_{i=1}^{n}\Delta x_i^4\right) + 9ma^4\sigma_q^4\sigma_w^4||\Delta x||^4
\end{aligned}$$

So since

$$Var[||\Delta y||^2] = \sum_{i=1,j=1,j\neq i}^{p} E[\Delta y_i^2 \Delta y_j^2] + \sum_{i=j=1}^{p} E[\Delta y_i^4] - p^2m^2a^4\sigma_q^4\sigma_w^4||\Delta x||^4$$