

## 5.2 連鎖律

計算グラフの順伝播は、計算結果を左から右へ伝達するものだった。一方、逆方向の伝播は「局所的な微分」を右から左へ伝達するものである。この「局所的な微分」を伝達する原理は、**連鎖律** (chain rule) によるものである。

### 5.2.1 計算グラフの逆伝播

早速、 $y = f(x)$  という計算があるとして、この計算の逆伝播を図 5-6 に表す。

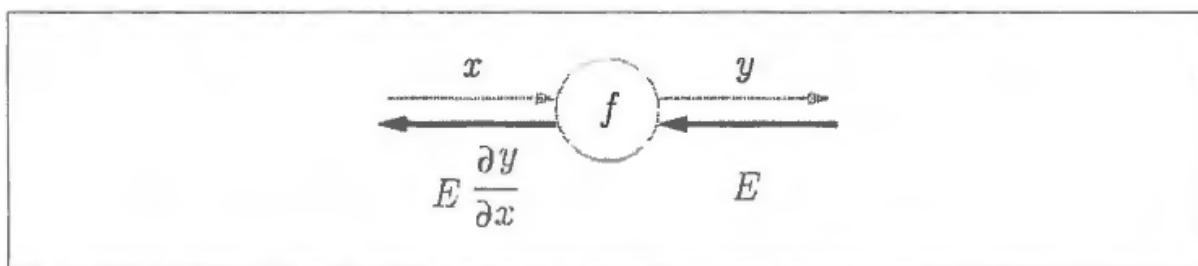


図 5-6 計算グラフの逆伝播：順方向とは逆向きに、局所的な微分を乗算する

図 5-6 に示すように、逆伝播の計算は信号  $E$  に対し、ノードの局所的な微分  $(\partial y / \partial x)$  を乗算し、それを次のノードに伝達していく。局所的な微分とは、 $y = f(x)$  の微分を求めるということである。このような計算手順を踏むことで、逆伝播は目的とする微分の値を効率よく求められる。これには、連鎖律が関係している。

### 5.2.2 連鎖律とは

連鎖律の説明には、まず**合成関数**を説明する必要がある。合成関数とは複数の関数によって構成される関数のことである。例えば、 $z = (x + y)^2$  という式は、次の式 (5.1) のように、2 つの式で構成される。

$$\begin{aligned} z &= t^2 \\ t &= x + y \end{aligned} \tag{5.1}$$

そして、連鎖律とは合成関数の微分が、合成関数を構成する各関数の微分の積によって表される性質のことである。式 (5.1) の例で言えば、 $\partial z / \partial x$  は  $\partial z / \partial t$  と  $\partial t / \partial x$  の積によって表すことができる。数式では、式 (5.2) のように書ける。

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial t} \frac{\partial t}{\partial x} \tag{5.2}$$

### 5.2.3 連鎖律と計算グラフ

では、式 (5.1) の連鎖律を用いた微分を、計算グラフで表してみる。2 乗の計算を「\*\*2」というノードで表すと、図 5-7 のように書ける。

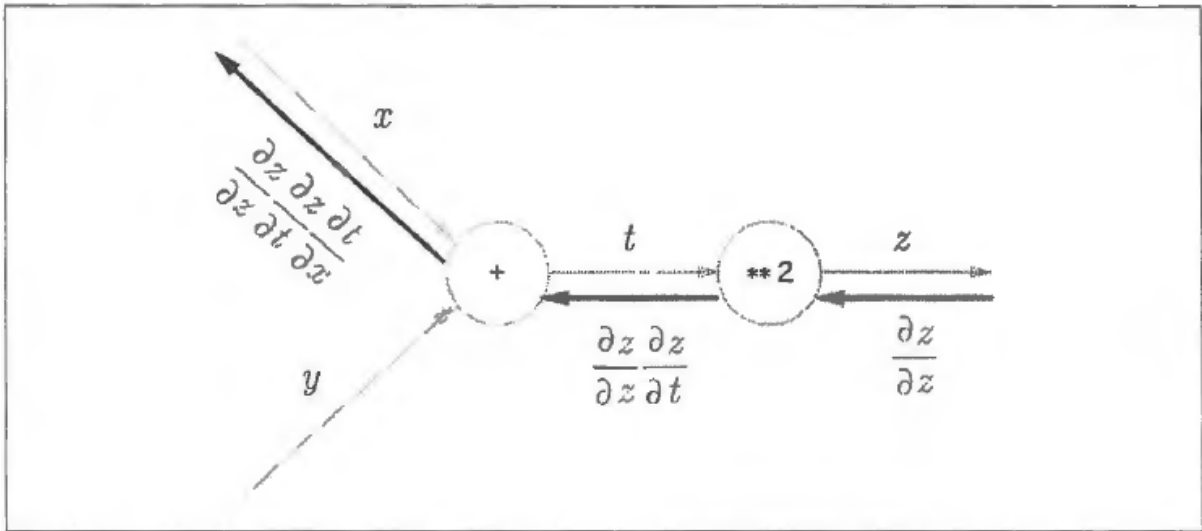


図 5-7 式 (5.4) の計算グラフ：順方向とは逆向きの方向に、局所的な微分を乗算して渡していく

図に示したように、計算グラフの逆伝播は右から左へと信号を伝播していく。図では、「\*\*2」への入力は  $\partial z / \partial z$  であり、これに局所的な微分である  $\partial z / \partial t$  を乗算して、次のノードへ渡していく。なお、 $\partial z / \partial z = 1$  であるので、数式上は省略できる。

図 5-7 で注目すべきは、一番左の逆伝播の結果である。これは、連鎖律より、

$$\frac{\partial z}{\partial z} \frac{\partial z}{\partial t} \frac{\partial t}{\partial x} = \frac{\partial z}{\partial t} \frac{\partial t}{\partial x} = \frac{\partial z}{\partial x}$$

が成り立ち、「 $x$  に関する  $z$  の微分」に対応する。つまり、逆伝播が行っていることは、連鎖律の原理から構成されている。図 5-7 に計算結果を代入すると、図 5-8 のようになる。

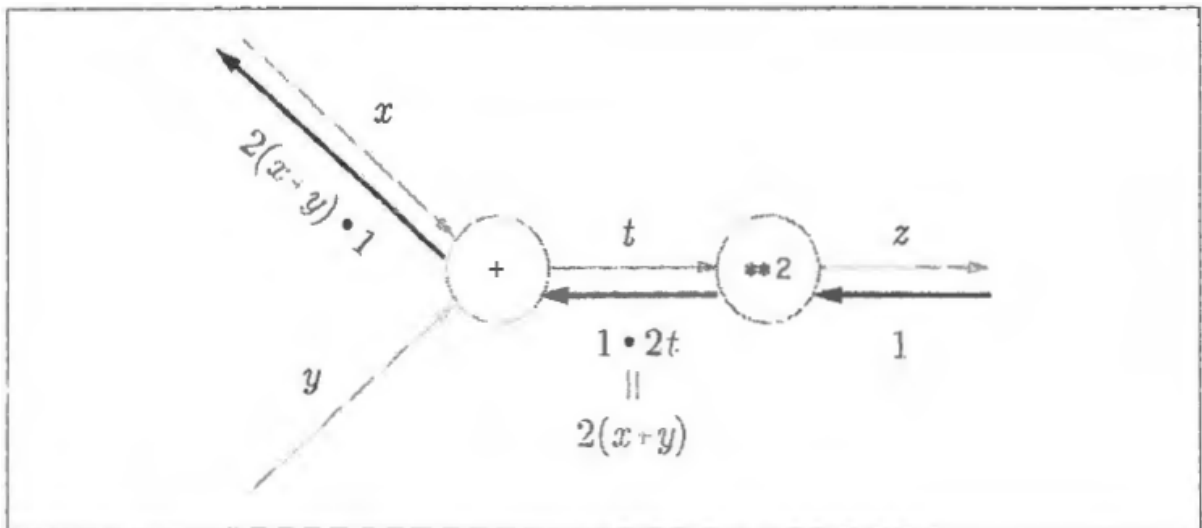


図 5-8 計算グラフの逆伝播の結果より、は  $2(x+y)$  となる