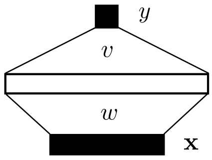

図 3.30 入力 $\mathbf{x}$ から出力 y を予測する, 1 層のニューラルネットワーク.

それでは, 入力 $\mathbf{x}$ に対する y の分布はどのようになるでしょうか. 式 (3.112) の和の中の各項 $v_j h_j(\mathbf{x})$ の v および w に関する期待値は,

$$\mathbb{E}_v[\mathbb{E}_w[v_j h_j(\mathbf{x})]] = \mathbb{E}_v[v_j]\mathbb{E}_w[h_j(\mathbf{x})] = 0 \tag{3.115}$$

です. また分散は, 上の事実から

$$\mathbb{V}_v[\mathbb{E}_w[v_j h_j(\mathbf{x})]] = \mathbb{V}_v[v_j \mathbb{E}_w[h_j(\mathbf{x})]] = \sigma_v^2 \mathbb{E}_w[h_j(\mathbf{x})]^2 \tag{3.116}$$

となります. すなわち, 確率変数 $v_j h_j(\mathbf{x})$ は j によらず平均 0 で, 有限の同じ分散を持つことになります.

ゆえに, 式 (3.112) のようにその H 個の和である y は, 中心極限定理によってガウス分布 $\mathcal{N}(0, H\sigma_v^2 \mathbb{E}_w[h_j(\mathbf{x})]^2)$ に近づきます. *27

このままでは, H が大きくなると y の値域も大きくなってしまうため, v_j の分布を

$$v_j \sim \mathcal{N}(0, \sigma_v^2/H) \tag{3.117}$$

と定義し直せば, $H \to \infty$ のとき $\lim_{h\to\infty} p(y) = \mathcal{N}(0, \sigma_v^2 \mathbb{E}_w[h_j(\mathbf{x})]^2)$ となります.

同様にして, 入力がそれぞれ $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ のとき, 対応する出力 $\mathbf{y} = (y_1, y_2, \ldots, y_N)$ は平均 0, 共分散

$$V[\mathbf{x}_n, \mathbf{x}_{n'}] = \sigma_v^2 \mathbb{E}[h(\mathbf{x}_n)h(\mathbf{x}_{n'})] \tag{3.118}$$

のガウス分布に近づきます. このことは任意の N および $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ について成り立ちますから, 本章の議論から, N 次元の $\mathbf{y}$ はガウス過程に従うことになります.