# 1 Symbol

## 1.1 Constant

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9$$

$$\alpha, \beta, \gamma, \delta, \epsilon(\varepsilon), \zeta, \eta, \theta(\vartheta), \iota, \kappa, \lambda, \mu, \nu, \xi, o, \pi, \rho(\varrho), \sigma, \tau, \upsilon, \pi(\varphi), \chi, \psi, \omega$$

## 1.2 Scalar

$$a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z$$

## 1.3 Vector

$$\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{f}, \mathbf{g}, \mathbf{h}, \mathbf{i}, \mathbf{j}, \mathbf{k}, \mathbf{l}, \mathbf{m}, \mathbf{n}, \mathbf{o}, \mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{s}, \mathbf{t}, \mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z}$$

## 1.4 Matrix

$$\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{E}, \mathbf{F}, \mathbf{G}, \mathbf{G}, \mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}, \mathbf{M}, \mathbf{N}, \mathbf{O}, \mathbf{P}, \mathbf{Q}, \mathbf{R}, \mathbf{S}, \mathbf{T}, \mathbf{U}, \mathbf{V}, \mathbf{W}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}$$

## 1.5 Tensor

$$\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}, \mathsf{E}, \mathsf{F}, \mathsf{G}, \mathsf{G}, \mathsf{I}, \mathsf{J}, \mathsf{K}, \mathsf{L}, \mathsf{M}, \mathsf{N}, \mathsf{O}, \mathsf{P}, \mathsf{Q}, \mathsf{R}, \mathsf{S}, \mathsf{T}, \mathsf{U}, \mathsf{V}, \mathsf{W}, \mathsf{X}, \mathsf{Y}, \mathsf{Z}$$

## 1.6 Set

$$\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}, \mathcal{G}, \mathcal{I}, \mathcal{J}, \mathcal{K}, \mathcal{L}, \mathcal{M}, \mathcal{N}, \mathcal{O}, \mathcal{P}, \mathcal{Q}, \mathcal{R}, \mathcal{S}, \mathcal{T}, \mathcal{U}, \mathcal{V}, \mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$$

# 2 Statistics

## 2.1 Probability

$$x \sim \mathcal{X} \tag{1}$$

$$x \doteq \mathcal{X} \tag{2}$$

$$x \leftarrow \mathcal{X} \tag{3}$$

$$p\left(\alpha\right), p\left(a\right), p\left(\mathbf{a}\right), p\left(\mathbf{A}\right), p\left(\mathsf{A}\right), p\left(\mathcal{A}\right)$$

$$p\left(\alpha \mid \beta\right), p\left(a \mid b\right), p\left(\mathbf{a} \mid \mathbf{b}\right), p\left(\mathbf{A} \mid \mathbf{B}\right), p\left(\mathsf{A} \mid \mathsf{B}\right), p\left(\mathcal{A} \mid \mathcal{B}\right)$$

# 3 Indexing

$$\underset{a \in \mathcal{A}}{\operatorname{argmax}} \; p\left(a\right)$$
$$\underset{x \in \mathcal{X}}{\operatorname{argmin}} \; p\left(x \mid y\right)$$

# 4 Distribution

$$\sigma\left(x\right)$$

$$\frac{\exp\left(p\left(x\right)\right)}{\sum_{x' \in \mathcal{X}} \exp\left(p\left(x'\right)\right)}$$

# 5 Neural Networks

## 5.1 Activation

$$\max\left(\mathbf{0}, \mathbf{x}\right)$$

$$\tanh\left(\mathbf{x}\right) \tag{4}$$

$$\{0, 1\}$$
$$\{a, \ldots, z\}$$

$$\mathbf{x} \oplus \mathbf{y}$$
$$\mathbf{x} \ominus \mathbf{y}$$
$$\mathbf{x} \odot \mathbf{y}$$
$$\mathbf{x} \oslash \mathbf{y}$$
$$\mathbf{x} \otimes \mathbf{y}$$

$$\mathbf{x} \|^{1} \mathbf{y} \tag{5}$$

$$\ell_\theta = \mathbb{E} - \log\left(p\left(y_t^* \mid \mathbf{h}_{t-1}, y_{t-1}; \theta\right)\right) \tag{6}$$

$$\mathbb{1}[x] \tag{7}$$