

# 1 Symbol

## 1.1 Constant

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

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

$$\mathbb{1}, \mathcal{N}, \mathcal{R}$$

## 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

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

## 1.6 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

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

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

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

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

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

### 3 Indexing

$$\operatorname{argmax}_{a \in \mathcal{A}} p(a)$$

$$\operatorname{argmin}_{x \in \mathcal{X}} p(x \mid y)$$

### 4 Distribution

$$\sigma(x)$$

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

### 5 Neural Networks

#### 5.1 Activation

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

$$\tanh(\mathbf{x})$$

$$\{0,1\}$$

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

$$\begin{array}{ll} \mathbf{x} \oplus \mathbf{y} & ( \ 26.85016\mathrm{pt} ) \\ \mathbf{x} \ominus \mathbf{y} & ( \ 26.85016\mathrm{pt} ) \\ \mathbf{x} \odot \mathbf{y} & ( \ 26.85016\mathrm{pt} ) \\ \mathbf{x} \oslash \mathbf{y} & ( \ 26.85016\mathrm{pt} ) \end{array}$$

$$x\,\overline{*}\,y\tag{1}$$

$$\mathbf{A}^\top$$

$$\mathbf{A}^{-1}$$

$$\mathbf{A}^{-\top}$$

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

$$\begin{aligned}\mathbf{i}_t &= \sigma(\mathbf{W}_i \cdot \mathbf{x}_t + \mathbf{b}_i) \\ \mathbf{f}_t &= \sigma\Big(\mathbf{W}_f \cdot \mathbf{x}_t + \mathbf{b}_f\Big) \\ \mathbf{o}_t &= \sigma(\mathbf{W}_o \cdot \mathbf{x}_t + \mathbf{b}_o) \\ \mathbf{g}_t &= \tanh(\mathbf{W}_g \cdot \mathbf{x}_t + \mathbf{b}_g) \\ \mathbf{c}_t &= \mathbf{i}_t \odot \mathbf{f}_t\end{aligned}$$

$$\begin{aligned} & -\log\Big(p(y_t \mid \mathbf{h}_{t-1}, y_{t-1}; \theta)\Big) \\ & \epsilon \sim \mathcal{N}\Big(0, (\sigma/t)^2\Big) \end{aligned}$$

$$\begin{aligned} & \max_{x \in \mathcal{X}} f(x) \\ & \min_{x \in \mathcal{X}} f(x) \\ & \sum_{x \in \mathcal{X}} f(x) \\ & \frac{1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} f(x) \end{aligned}$$

$$\{f(x) \, | \, x \in \mathcal{X}\}$$

$$\begin{array}{c} \frac{\partial f(x)}{\partial x} \\ \frac{\partial^2 f(x)}{\partial x^2} \\ \frac{\partial^2 f(x)}{\partial x \, \partial y} \end{array}$$

$$\mathbb{E}_{s_t \sim E, a_i \sim \pi} \big[ R(s_t, a_t) \big] \tag{2}$$

$$\mathbf{y} = \mathbf{W}_l^{[1,3]} \ast \mathbf{x} + \mathbf{b}_l \tag{3}$$

## 6 Reinforcement Learning

$$\nabla_{\mathbf{x}} f(\mathbf{x})$$

$$\nabla_{\mathbf{x}}^2 f(\mathbf{x})$$