引入.

- PRML 第一章两个地方引入变分法,

Loss function for 回归

定义回归损失

- $E[L] = \iint L(t, y(x))\, p(x, t)\, dx\, dt.$

通常我们选择平方差即.

$$L(t, y(x)) = [y(x) - t]^2$$

此时代价为

$$E[L] = \iint [y(x) - t]^2 p(x, t)\, dx\, dt$$

求回归问题的目标就是选择一个函数 $y(x)$ 来最小化代价函数，通过变分方法，我们可以求解一个普通的 $y(x)$ 来满足这个要求。方法如下：

这里是邻域.

设 $f(x) = y(x) + \varepsilon \eta(x)$，定义泛函 $\phi(\varepsilon) = F[f] = F[y + \varepsilon\eta]$.

$\varepsilon \to 0$

那么 $\phi(0) = F[y] = E[L]$

给定在邻域内.

$$\phi(\varepsilon) = \iint [y(x) + \varepsilon \eta(x) - t]^2 p(x, t)\, dx\, dt.$$

ε控制邻域大小.

若 y 为 $F[f]$ 的极值点，则 $\phi'(0) = 0$，即 $\frac{d\phi}{d\varepsilon}\Big|_{\varepsilon=0} = 0$，其中

$$\frac{d\phi}{d\varepsilon} = 2\iint \eta\,(y + \varepsilon\eta - t)\, p(x, t)\, dx\, dt.$$

带入 $\varepsilon = 0$

$$\frac{d\phi}{d\varepsilon}\Big|_{\varepsilon=0} = 2\iint \eta\,(y - t)\, p(x, t)\, dx\, dt.$$

$$= 2\int \eta \left\{ \int (y - t)\, p(x, t)\, dt \right\} dx$$

若对任意 η，都有 $\phi'(0) = 0$

则 $\int (y - t)\, p(x, t)\, dt = 0.$

$$\Rightarrow y(x) = \frac{\int t\, p(x, t)\, dt}{p(x)} = \int t\, p(t|x)\, dt = E_t[t|x].$$

信息论

P.53. • 推导当连续随机变量 x 的熵最大时，x 服从高斯分布.

公式 •
$$H[x] = -\int p(x)\ln p(x)\,dx$$
$$\int_{-\infty}^{+\infty} p(x)\,dx = 1$$
$$\int_{-\infty}^{+\infty} x\,p(x)\,dx = \mu$$
$$\int_{-\infty}^{+\infty} (x-\mu)^2 p(x)\,dx = \sigma^2$$

利用拉格朗日. •
$$L = -\int_{-\infty}^{+\infty} p(x)\ln p(x)\,dx + \lambda_1\left(\int_{-\infty}^{+\infty} p(x)\,dx - 1\right) + \lambda_2\left(\int_{-\infty}^{+\infty} x\,p(x)\,dx - \mu\right) + \lambda_3\left(\int_{-\infty}^{+\infty}(x-\mu)^2 p(x)\,dx - \sigma^2\right)$$

变分法求解 • 设 $f(x) = p(x) + \varepsilon\eta(x)$，代入 L 中有 $\phi(\varepsilon) = L[f]$.

若 $p(x)$ 能取极值，$\phi'(0) \equiv 0$.
$$\phi(\varepsilon) = -\int (p+\varepsilon\eta)\ln(p+\varepsilon\eta)\,dx + \lambda_1\left[\int (p+\varepsilon\eta)\,dx - 1\right] + \lambda_2\left[\int x(p+\varepsilon\eta) - \mu\right] + \lambda_3\left[\int (x-\mu)^2(p+\varepsilon\eta) - \sigma^2\right].$$
$$\phi'(\varepsilon) = -\left[\int \eta\ln(p+\varepsilon\eta)\,dx + \int \eta\,dx\right] + \lambda_1\left[\int \eta\,dx\right] + \lambda_2\left[\int x\,\eta\,dx\right] + \lambda_3\left[\int (x-\mu)^2\cdot\eta\,dx\right]$$

当 $\varepsilon = 0$ 时.
$$\phi'(0) = -\left[\int \eta\ln p\,dx + \int \eta\,dx\right] + \lambda_1\int\eta\,dx + \lambda_2\int x\,\eta\,dx + \lambda_3\int (x-\mu)^2\eta\,dx$$
$$= \int \eta\left[-\ln p - 1 + \lambda_1 + \lambda_2 x + \lambda_3 (x-\mu)^2\right]dx$$

若对任意 η 都有 $\phi'(0)=0$. 则 $\ln p = -1 + \lambda_1 + \lambda_2 x + \lambda_3 (x-\mu)^2$

即 $p(x) = \exp[-1 + \lambda_1 + \lambda_2 x + \lambda_3 (x-\mu)^2]$.