$\mathcal{N}(x|\mu,\tau)$ $\quad$ $p(\mu,\tau)\cdot p(D|\mu,\tau) \propto p(\mu,\tau|D)$

$$p(\mu,\tau|a_0,b_0,\mu_0,\lambda_0) = \mathrm{Gamma}(\tau|a_0,b_0)\cdot \mathcal{N}(\mu|\mu_0,\lambda_0\tau)$$

$$\ln p(\mu,\tau|x_1 .. x_N) = \text{const} + (a_0-1)\ln\tau - b_0\tau + \frac{1}{2}\ln\tau - \frac{\lambda_0\tau}{2}(\mu-\mu_0)^2 -$$

$$+\frac{N}{2}\ln\tau - \frac{\tau}{2}\sum_{n=1}^{N}(x_n-\mu)^2 \quad (N\mu^2 - 2\mu\sum x_n + \sum x_n^2)$$

$\frac{\lambda_N\tau}{2}(\mu-\mu_N)^2$

$\tau\frac{\lambda_N\mu_N^2}{2}$

$$\mu^2: \ -\frac{1}{2}\lambda_0\tau - \frac{1}{2}\tau\cdot N$$

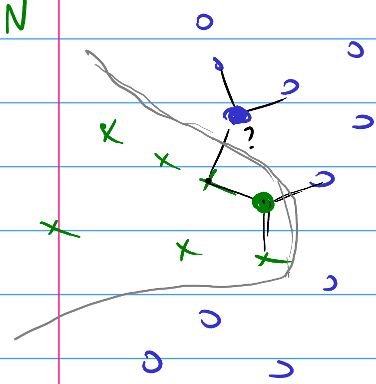
$$\mu: \ \mu_0\lambda_0\tau + \tau\sum x_n = \lambda_N\tau\left(\frac{\mu_0\lambda_0+\sum x_n}{\lambda_N}\right)$$

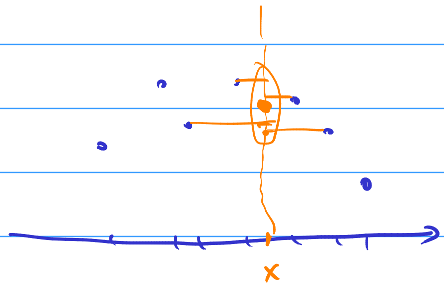
$$\lambda_N = \lambda_0 + N$$
$$\mu_N = \frac{\mu_0\lambda_0 + \sum x_n}{\lambda_0 + N}$$

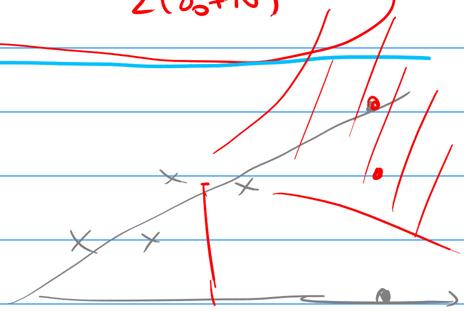
$$a_N = a_0 + \frac{N}{2}$$

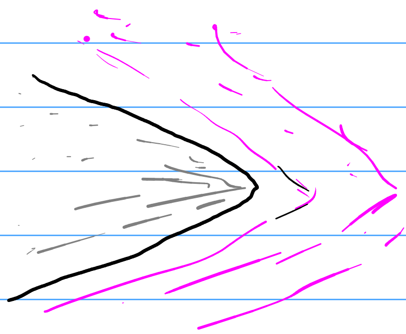
$$b_N = b_0 + \frac{\lambda_0\mu_0^2}{2} + \frac{1}{2}\sum x_n^2 - \frac{(\mu_0\lambda_0+\sum x_n)^2}{2(\lambda_0+N)}$$

kNN



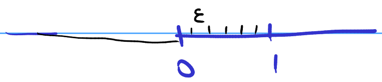




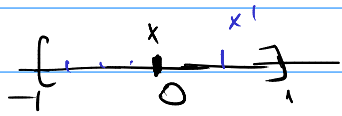


Curse of dimensionality

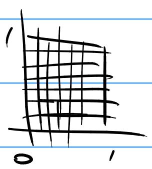
1) $\mathbb{E}_{\dots} = \int f(\bar{x})\,p(\bar{x})\,d\bar{x} \approx$ $\quad$ 2) $E[\|\bar{x}-\bar{x}'\|^2] = ?$



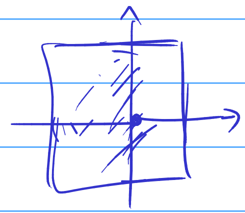
$1/\varepsilon$



$$\mathbb{E}[x'^2] = x^3/3\Big|_0^1 = 1/3$$



$1/\varepsilon^2$

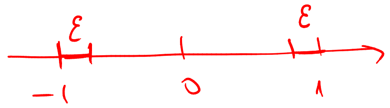


$$d(x',0)^2 = \mathbb{E}[x^2+y^2] = 2/3$$

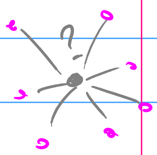
d $\quad$ $1/\varepsilon^d$ $\quad$ d

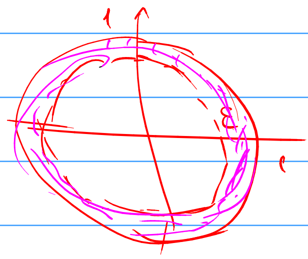
$$\bar{X} = (x_1 \dots x_{1000})$$

3)



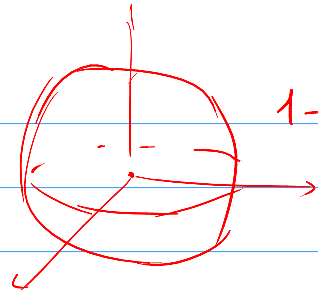
ε





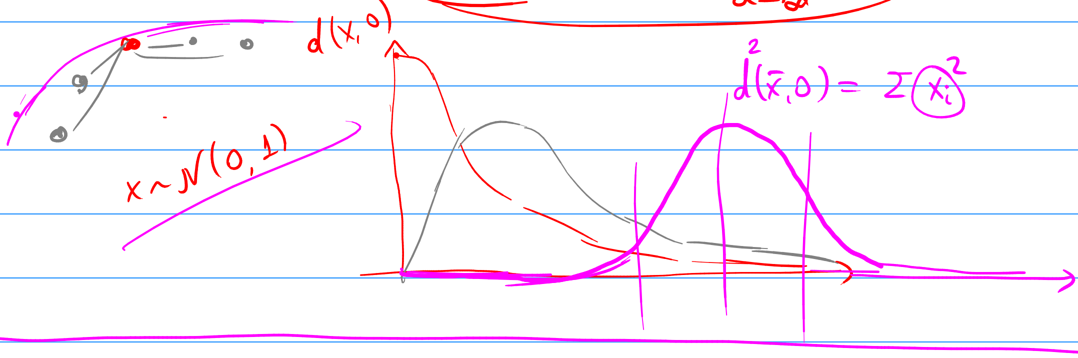
$$\frac{\pi\cdot 1^2-\pi\cdot(1-\varepsilon)^2}{\pi\cdot 1^2}$$

$$1-(1-\varepsilon)^2$$



$$1-(1-\varepsilon)^3$$

$$1-(1-\varepsilon)^d \xrightarrow[d\to\infty]{} 1$$



$$d^2(\bar{x},0)=\sum x_i^2$$

$\bar{x}\in\mathbb{R}^p \quad f(\bar{x}) \sim y \qquad p(\bar{x},y)=p(\bar{x})\,p(y|\bar{x})$

$f(\bar{x})$

$$L(y,f(\bar{x}))=(y-f(\bar{x}))^2$$

$$EPE[f]=\mathbb{E}_p\left[L(f(\bar{x}),y)\right]=\int (y-f(\bar{x}))^2\,p(\bar{x},y)\,d\bar{x}\,dy \xrightarrow[f]{} \min$$

$$\int(y-a)^2 p(y)\,dy \xrightarrow[a]{} \min \qquad \int\left[\int (y-f(\bar{x}))^2\,p(y|\bar{x})\,dy\right]p(\bar{x})\,d\bar{x}$$

$$a^*=E_p[y] \qquad f(\bar{x}) \quad \min$$

$$\int(y-Ey+Ey-a)^2\,p(y)\,dy=$$

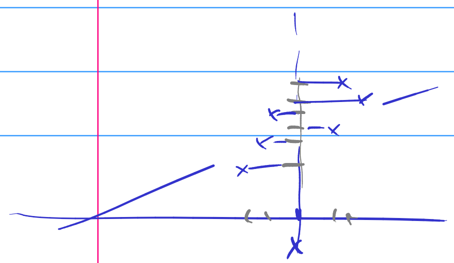
$$=\int(y-Ey)^2 p(y)\,dy+\int(\ldots-\ldots)^2-2\int\cancel{(y-Ey)}(a-Ey)\,p(y)\,dy$$

$$\hat{f}(\bar{x})=\mathbb{E}_{p(y|\bar{x})}[y]=\int y\,p(y|\bar{x})\,dy$$

regression function

$$\hat{f}(\bar{x})=\mathbb{E}_{y|\bar{x}}[y]\approx\frac{1}{k}\sum_{i=1}^{k}y_i \approx \frac{1}{k}\sum_{y_i\in kNN(\bar{x})}y_i$$

$y_i \sim p(y|\bar{x})$



$$EPE[f] = \int (\underbrace{y - f(\bar{x})}_{y - \hat{f} + \hat{f} - f})^2 \, p(\bar{x}, y) \, d\bar{x} \, dy =$$

$$= \int (y - \hat{f}(\bar{x}))^2 \ldots + \underbrace{2 \int (y - \hat{f}(\bar{x}))(\hat{f}(\bar{x}) - f(\bar{x})) \, p(\bar{x}, y) \, d\bar{x} \, dy}_{} + \int (\hat{f} - f)^2 \, p(\bar{x}, y) \, d\bar{x} \, dy =$$

$$+2 \int \Big[\underbrace{\int (y - \hat{f}(\bar{x})) \, p(y|\bar{x}) \, dy}_{=0} \Big] (\hat{f}(\bar{x}) - f(\bar{x})) \, p(\bar{x}) \, d\bar{x}$$

$$EPE[f] = \underbrace{\mathbb{E}[(y - \hat{f}(\bar{x}))^2]}_{\text{NOISE}} + \underline{\mathbb{E}[(\hat{f}(\bar{x}) - f(\bar{x}))^2]}$$

$$f(\bar{x}; D) \qquad D \sim p(\bar{x}, y)$$

$$\mathbb{E}[(\hat{f} - f)^2] = \mathbb{E}[(\hat{f} - \mathbb{E}_D f + \mathbb{E}_D f - f)^2] =$$

$$= \underline{\int (\hat{f} - \mathbb{E}_D f)^2 \ldots} + \cancel{2 \int (\hat{f} - \mathbb{E}_D f)(\mathbb{E}_D f - f) \ldots} + \underline{\int (\mathbb{E}_D f - f)^2 \ldots}$$

$$\underset{\text{ERROR}}{EPE[f]} = \underset{\text{NOISE}}{\mathbb{E}[(y - \hat{f}(\bar{x}))^2]} + \underset{\text{BIAS}^2}{\mathbb{E}[(\hat{f}(\bar{x}) - \mathbb{E}_D f(\bar{x}, D))^2]} + \underset{\text{VARIANCE}}{\mathbb{E}[(f(\bar{x}) - \mathbb{E}_D f(\bar{x}, D))^2]}$$

ERROR = NOISE + BIAS² + VARIANCE

Bias-variance-noise decomposition