References

1. Pattern Classification

Duda, Hart, Stork

Under the $\{0,1\}$ loss function

Bayes Classifier is

$$i^*(x) = \underset{i \in \{1,\dots,M\}}{\operatorname{argmax}} \eta_i(x)$$

$$= \underset{i \in \{1,\dots,m\}}{\operatorname{argmax}} f(\eta_i(x))$$

$f : [0,1] \to \mathbb{R}$ is an increasing fn.

$f(t) \leq f(s)$ whenever $t < s$

$$\eta_i(x) = P(Y = i \mid X = x)$$

$$P(Y = i \mid X = x) = N(x \mid \mu_i, C_i)$$

$$\log \eta_i(x) = \log P(Y = i \mid X = x)$$

$$= \log \frac{p_i}{(\sqrt{2\pi})^d |C_i|^{1/2}} e^{-\frac{1}{2}(x-\mu_i)^T C_i^{-1} (x-\mu_i)}$$

$$g_i(x) = \log p_i - \frac{1}{2} \log |C_i|$$

$$- \frac{1}{2}(x-\mu_i)^T C_i^{-1} (x-\mu_i)$$

$$i^*(x) = \underset{i}{\operatorname{argmax}} \; g_i(x)$$

$$\min_{t \in \mathbb{R}} f(t)$$

Find t^* $\quad f(t^*) \le f(t)$

$$f(t) = \frac{1}{2}at^2 + bt + c \qquad a > 0$$

$$= \frac{1}{2}a\left(t + \frac{b}{a}\right)^2 + c - \frac{b^2}{2a}.$$

$$f(t) \geq f(t^*) = c - \frac{b^2}{2a}.$$

$$t^* = -\frac{b}{a}.$$

$$f(t) = \frac{1}{2}\|u - tv\|^2 \qquad u, v \in \mathbb{R}^d$$

$f(t) \geq 0$ for all $t \in \mathbb{R}^d$

$$\|x\| = \sqrt{x^T x} = \sqrt{\sum_{i=1}^{d} x_i^2}$$

$$\|x\| = 0 \Rightarrow \sum_{i=1}^{d} x_i^2 = 0$$

$$\Rightarrow \quad x_i = 0 \qquad i \in \{1, \dots, d\}$$

$$f(t) = \frac{1}{2}\|u\|^2 - t u^T v + \frac{t^2}{2}\|v\|^2$$

$$t^* = \frac{(u^T v)}{\|v\|^2}$$

$$f(t^*) = \frac{1}{2}\|u\|^2 - \frac{(u^T v)^2}{2\|v\|^2} \geq 0$$

$$f'(t^*) = (u - t^* v)^T v = 0$$

For any $u, v \in \mathbb{R}^d$

$$\|u\| \|v\| \geq |u^T v|$$

Equality holds only when $u = \alpha v$ for some $\alpha \in \mathbb{R}$

$$\max_{x \in \mathbb{R}^d} x^T A x$$

$A \in S_d^+$ $\quad A \in \mathbb{R}^{d \times d}$

$A = A^T$

$x^T A x > 0$

$u = x, \quad v \doteq Ax$

$$u^T v \leq \|u\| \|v\|$$

Equality iff
$Ax = t\, x$

$$x^T A x \leq t \|x\|^2$$

$\lambda_{max}(A)$ is the largest eigenvalue of A

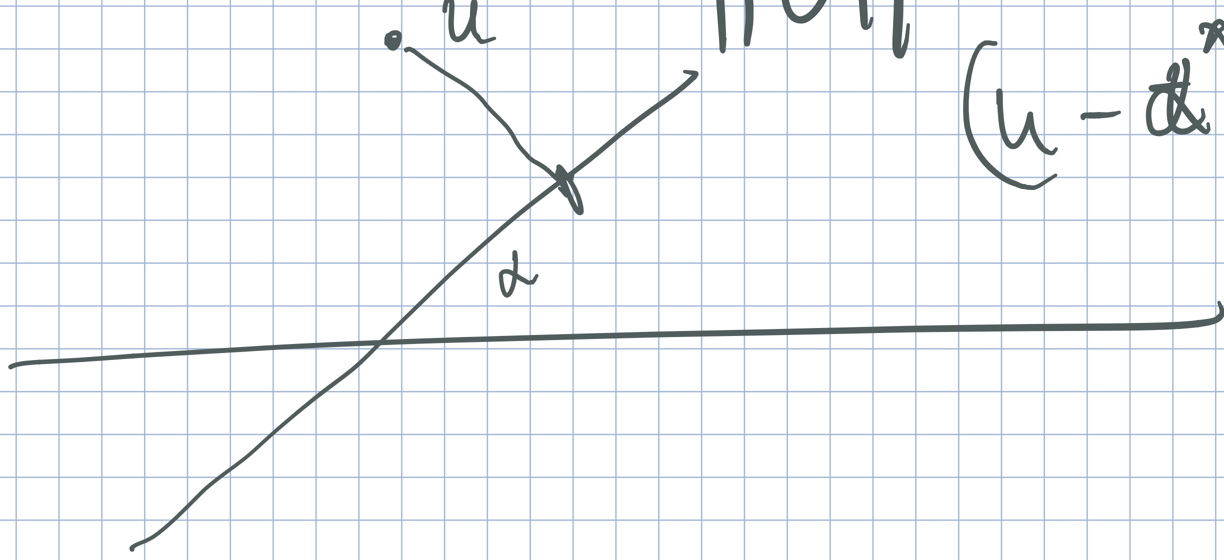
$$t = \lambda_{max}(A) \quad \max_{x \setminus \{0\}} \frac{x^T A x}{x^T x} = \lambda_{max}(A)$$

$C = \{\alpha v \mid \alpha \in \mathbb{R}\}$

Projection of any $u \in \mathbb{R}^d$ on C is the point in C which is closest u

$$\min_{\alpha} \frac{1}{2} \| u - \alpha v \|^2$$

$$\Rightarrow \alpha^* = \frac{u^T v}{\|v\|}$$



$(u - \alpha^* v)^T v = 0$

For every u and v there exists an unique α

$$u = \alpha v + v_\perp$$

$$v_\perp^T v = 0 \qquad v_\perp = u - \alpha v$$

$$\max_{x \in \mathbb{R}^d} x^T A x \qquad A \in S_d^+ \qquad A \in \mathbb{R}^{d \times d}$$

$$A = A^T$$

$u = x$, $\quad v \doteq Ax$

$$x^T A x > 0$$

$$u^T v \leq \|u\| \|v\|$$

Equality iff $\qquad \lambda_{max}(A)$ is the largest eigenvalue of A

$$Ax = t\, x$$

$$x^T A x \leq t \|x\|^2$$

$$t = \lambda_{max}(A) \qquad \max_{x \setminus \{0\}} \frac{x^T A x}{x^T x} = \lambda_{max}(A)$$

Let $A \in S_d^{+}$ and $B \in S_d^{++}$

$$\max_{x \in \mathbb{R}^d \setminus \{0\}} \frac{x^T A x}{x^T B x}$$

There exists E

$B = E^2$ $\quad E \in S_d$, full rank

$(E = B^{1/2})$

$y = E x$, $\quad x = E^{-1} y$

$$\max_{y \neq 0} \frac{y^T E^{-1} A E^{-1} y}{y^T y} = \lambda_{max}(E^{-1} A E^{-1})$$

$$\therefore \quad E^{-1} A E^{-1} y = \lambda_{max}(E^{-1} A E^{-1})\, y$$

$$E^{-1} A x = \lambda_{max}(E^{-1} A E^{-1})\, E x$$

$$A x = \lambda_{max}(E^{-1} A E^{-1})\, B x$$

$Ax = \lambda B x$ is generalized eigenvalue problem