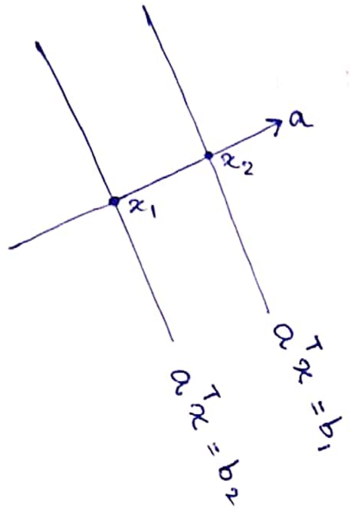
Sagnik Ghosh Assignment II 3343-6044

1. $a^Tx = b_1$ — (i) $a^Tx = b_2$ — (ii)



The edualions are of hyperplanes, we know that any two points belonging to these two hyperplanes must be on the same normal vector which is perpendicular to both these planes.

Let x_1 be a point on 1st plane, $a^Tx_1 = b_1$
x_2 be a point on 2nd plane, $a^Tx_2 = b_2$

Now x_1, x_2 is scalar multiple of vector a, as these are on the line of projection as of a.

$$x_1 = \alpha_1 a \qquad x_2 = \alpha_2 a$$

now $a^T(\alpha_1 a) = b_1$ $\qquad a^T(\alpha_2 a) = b_2$

$$\therefore \alpha_1 = \frac{b_1}{\|a\|^2} \qquad \therefore \alpha_2 = \frac{b_2}{\|a\|^2}$$

$$\|x_1 - x_2\| = \|\alpha_1 a - \alpha_2 a\| = \left\| \frac{b_1}{\|a\|^2} a - \frac{b_2 a}{\|a\|^2} \right\| = \frac{|b_1 - b_2| \|a\|}{\|a\|^2}$$

$$= |b_1 - b_2| / \|a\|.$$

2. a) $E[x] = a_1p_1 + a_2p_2 + \dots + a_Kp_K = a^Tp$

a^Tp is a linear function hence this is both concave and convex.

b) $Pr[x > \alpha] = p_{\alpha+1} + p_{\alpha+2} + \dots + p_K$

now we can have a vector s, where $s[0:\alpha] = 0$ and $s[\alpha+1:K] = 1$.

so, $Pr[x > \alpha] = s^Tp$ whic is linear, hence both concave and convex

c) $Pr[\alpha < x < \beta] = p_{\alpha+1} + p_{\alpha+2} + \dots + p_\beta$

similarly we can have a vector s, where $s[0:\alpha] = 0$ and $s[\alpha+1:\beta] = 1$, $s[\beta+1, K] = 0$

so, $Pr[\alpha < x < \beta] = s^Tp$, which is linear hence both concave and convex.

d) $-\sum_{i=1}^{k} p_i \log p_i$

we know negative entropy is a convex function ($p \log p$)

Further, additive property preserve convexity.

Therefore $\sum_{i=1}^{k} p_i \log p_i$ is also convex.

Now, if $f(x)$ is convex the $-f(x)$ is concave.

Hence $-\sum_{i=1}^{k} p_i \log p_i$ is concave function.

e) $Var_p(x) = E_p[X^2] - E_p[X]^2$

Expectation is a linear function hence, $E_{\lambda p+(1-\lambda)q}[x^2] = \lambda E_p[x^2] + (1-\lambda) E_q[x^2]$ — (I)

$E_p[x] = f(x)$ is linear, but $E_p[x]^2$ is convex in nature.

hence, we can apply Jensen's inequality ($f(E[z]) \le E[f(z)]$).

$$\lambda E_p[x]^2 + (1-\lambda) E_q[x]^2 \ge (\lambda E_p[x] + (1-\lambda) E_q[x])^2 = E_{\lambda p+(1-\lambda)q}[x]^2 \quad \text{— (II)}$$

Now, $Var_{\lambda p+(1-\lambda)q}(x) = E_{\lambda p+(1-\lambda)q}[x^2] - E_{\lambda p+(1-\lambda)q}[x]^2$

now substituting (I) and (II)

$$Var_{\lambda p+(1-\lambda)q}(x) \ge \lambda E_p[x^2] + (1-\lambda) E_q[x^2] - \lambda E_p[x]^2 - (1-\lambda) E_q[x]^2$$

$$\ge \lambda\, var_p(x) + (1-\lambda) Var_q(x)$$

Hence variance is a ~~convex~~ ~~fn~~ concave function.

3. minimize$_{\theta,u,v}$ $\sum_{i=1}^{m} (u_i^2 + 2Mv_i)$ subject to $-u - v \le \phi\theta - \psi \le u + v$ — (I)

$0 \le u \le M \cdot 1 \quad v \ge 0$

(I) can be rewritten as $|\phi_i^T\theta - \psi_i| \le u_i + v_i$

now, we can minimize u_i and v_i but both has to be greater than 0 from inequality condition.

From the other inequality, the lowest $u_i + v_i$ can achieve $|\phi_i^T\theta - \psi_i|$.

So, to minimize, $v_i = |\phi_i^T\theta - \psi_i| - u_i$ — (II)

now let's take $|\phi_i^T\theta - \psi_i| = |t_i|$ — (III)

now, the minimizing expression can be rewritten only in terms of u_i,

$$\underset{\theta}{\text{minimize}} \ \tilde{L}(\theta) \quad \text{subject to} \quad |\phi^T\theta - \psi_i| \le u_i + v_i$$

$$\tilde{L}(\theta) = \inf_u \sum_{i=1}^{m} (u_i^2 + 2M(|t_i| - u_i)) \quad \left[\begin{array}{l}\text{From (II)} \\ \text{and (III)}\end{array}\right]$$

$$= \inf_u \sum_{i=1}^{m} (u_i^2 - 2Mu_i + 2M|t_i|)$$

now, $u_i = \begin{cases} |t_i| & \text{when } |t_i| \le M \\ M & |t_i| > M. \end{cases}$

So, $\tilde{L}(\theta) = \begin{cases} t^2 & |t| \le M \\ M(2|t| - M) & |t| > M. \end{cases}$

$$\underset{\theta}{\text{minimize}} \sum_{i=1}^{m} h(t_i)$$