3a) GE 2014

~~p(y;~~

$$p(y;\phi) = (1-\phi)^{y-1}\phi$$

$$= \exp\left((y-1)\cdot\log(1-\phi) + \log(\phi)\right)$$

$$= \exp\left(y\cdot\log(1-\phi) - \log(1-\phi) + \log(\phi)\right)$$

$$= \exp\left(\log(1-\phi)\cdot y + \log\left(\frac{\phi}{1-\phi}\right)\right)$$

$\Rightarrow T(y) = y \qquad \eta = \log(1-\phi) \Rightarrow e^{\eta} = 1-\phi$

$\phi = 1-e^{\eta}$

$$a(\eta) = -\log\left(\frac{\phi}{1-\phi}\right) = -\log\left(\frac{1-e^{\eta}}{1-1+e^{\eta}}\right) = -\log\left(\frac{1-e^{\eta}}{e^{\eta}}\right)$$

$$= -\log\left(\frac{1}{e^{\eta}} - 1\right)$$

$$= -\log(1-e^{\eta}) + \eta$$

$b(y) = 1$

PE 2014

~~b)~~ ii)

$\cancel{h_\theta(x) = E[y|x;\theta]}$

$p(y;\eta) = (1-\phi)^{y-1}\phi$

$\phi = 1 - e^\eta$

$\Rightarrow p(y;\eta) = (1-1+e^\eta)^{y-1} \cdot (1-e^\eta)$

$= e^{\eta\cdot(y-1)} \cdot (1-e^\eta) = e^\eta$

$\Rightarrow \log(p(y;\eta)) = \eta \cdot (y-1) + \log(1-e^\eta)$

$\Rightarrow \log \prod_{i=1}^m p(y^{(i)}|x^{(i)};\theta) = \sum_{i=1}^m x^{(i)T}\theta\,(y^{(i)}-1) + \log(1-e^{x^{(i)T}\theta})$

7b) PE 2014 1/2 &

Classifier rule:

$$p(x|y=1)\cdot p(y=1) - p(x|y=0)\cdot p(y=0) \geq 0$$

or with logs:

$$\log\left(\frac{1}{(2\pi)^{n/2}} \cdot \frac{1}{|\Sigma_1|^{\frac{1}{2}}} \cdot \exp(\ldots)\right) + \log(\phi) - \log(\ldots) - \log(1-\phi)$$

$$= \log\left(\frac{1}{\pi \cdot \sqrt{\Sigma_1}}\right) - \frac{1}{2}(x-\mu_1)^T \Sigma_1^{-1}(x-\mu_1) + \log(\phi)$$

$$- \log\left(\frac{1}{\pi \cdot \sqrt{\Sigma_0}}\right) + \frac{1}{2}(x-\mu_0)^T \Sigma_1^{-1}(x-\mu_0) - \log(1-\phi)$$

$$= \frac{1}{2}\left((x-\mu_0)^T \Sigma_0^{-1}(x-\mu_0) - (x-\mu_1)^T \Sigma_1^{-1}(x-\mu_1)\right)$$

$$\underbrace{+ \log\left(\frac{\phi}{1-\phi}\right) + \log\left(\frac{\sqrt{|\Sigma_0|}}{\sqrt{|\Sigma_1|}}\right)}_{c}$$

linear part of params: PE 2014 2/2

$$x^T\Sigma_0^{-1}x - x^T\Sigma_0^{-1}\mu_0 - \mu_0^T\Sigma_0^{-1}x + \underbrace{\mu_0^T\Sigma_0^{-1}\mu_0}_{c}$$
$$- x^T\Sigma_1^{-1}x + x^T\Sigma_1^{-1}\mu_1 + \mu_1^T\Sigma_1^{-1}x - \underbrace{\mu_1^T\Sigma_1^{-1}\mu_1}_{c}$$

$$= x^T\Sigma_0^{-1}x - x^T\Sigma_1^{-1}x - 2(\mu_0^T\Sigma_0^{-1}x - \mu_1^T\Sigma_1^{-1}x)$$

$$= x^T(\Sigma_0^{-1}-\Sigma_1^{-1})x + 2\cdot(\mu_1^T\Sigma_1^{-1}-\mu_0^T\Sigma_0^{-1})\cdot x$$

(remember the $\frac{1}{2}$ from before)

$\Rightarrow A: \frac{1}{2}(\Sigma_0^{-1}-\Sigma_1^{-1})$

$B: \cancel{2}\cdot(\Sigma_1^{-1}\mu_1 - \Sigma_0^{-1}\mu_0)$

$C: \log\left(\frac{\phi}{1-\phi}\right) + \log\left(\frac{\sqrt{|\Sigma_0|}}{\sqrt{|\Sigma_1|}}\right) + \frac{1}{2}(\mu_0^T\Sigma_0^{-1}\mu_0 - \mu_1^T\Sigma_1^{-1}\mu_1)$

PE 2014

2a) Naive Bayes Classifier:

Show: $NB = \sum_{i=1} w_i c_i + b \geq 0$; $c =$ count of word i in document

Classify 1 if $p(y=1 \mid x) > p(y=0 \mid x)$

$\Rightarrow$ if $p(y=1 \mid x) - p(y=0 \mid x) > 0$

1 word

$\Rightarrow p(x \mid y=1)\, p(y=1) - p(x \mid y=0)\, p(y=0) > 0$

$\Rightarrow \sum_{j=1}^{n} \log(p(x_j \mid y=1)) + [\log(p(y=1)) - \log(p(x_j \mid y=0)) - \log(p(y=0))]$

$\Rightarrow b = \log\left(\frac{\phi_y}{1-\phi_y}\right)$

$\sum_{j=1}^{n} \log(p(x_j \mid y=1)) - \log(p(x_j \mid y=0))$

$= \sum_{j=1}^{n} \log(\phi_{x_j \mid y=1}) - \log(\phi_{x_j \mid y=0})$

$\Rightarrow w_i = \log\left(\frac{\phi_{x_i \mid y=1}}{\phi_{x_i \mid y=0}}\right)$

PE 2014 1/2 8

1) Prove:

$v^T v = 1$

$v v^T v = v$

$(vv^T y)^T = y^T (vv^T)^T = y^T (vv^T)$

$\tilde{J}(\hat{\theta}_{new}) \leq J(\hat{\theta})$

=> Show:

$$2 \cdot (XX^T y)^T (vv^T y) - 2 \cdot y^T (vv^T y) + (vv^T y)^T (vv^T y) \leq 0$$

~~Know: $\theta = X^T y$~~

$$\cancel{2 (vv^T y)^T \left(XX^T y - y + \frac{vv^T y}{2} \right)}$$

$$\cancel{\left((XX^T + vv^T) y - y \right)}$$

$$\cancel{\tilde{X}\tilde{X}^T y - \bar{y}}$$

$$\cancel{2 (XX^T y)^T (vv^T y) + (vv^T y)^T (vv^T y) \leq 2 y^T (vv^T y)}$$

$$\cancel{2 \cdot (vv^T y)^T (XX^T y - y) + (vv^T y)^T (vv^T y)}$$

$|b|_2 =$ PE 2014 ☆

$$(XX^T\vec{y})^T(XX^T\vec{y}) + (XX^T\vec{y})^T vv^T\vec{y} - (XX^T\vec{y})^T y$$

$$+ (vv^T\vec{y})^T(XX^T\vec{y}) + (vv^T\vec{y})^T(vv^T\vec{y}) - (vv^Ty)^T\vec{y}$$

$$- \vec{y}^T(XX^Ty) - y^T(vv^T\vec{y}) + y^Ty$$

$$= 3(0) + 2\cdot(XX^Ty)^T vv^T\vec{y} - 2\cdot y^T(vv^Ty) + (vv^Ty)^2$$

PE 2014 2/2

$$2 \cdot y^T X X^T v v^T y - 2 \cdot y^T v v^T y + y^T v \underline{v^T v} v^T y \leq 0$$

$$2 \cdot y^T X X^T v v^T y - 2 \cdot y^T v v^T y + y^T v v^T y \leq 0$$

$$2 \cdot y^T X X^T v v^T y - y^T v v^T y \leq 0$$

$$\cancel{(X X^T y)^T (v v^T y)} - (v^T y)^T (v^T y)$$

$\Rightarrow = 0$ because of orthonormality

$$\Rightarrow -y^T v v^T y \leq 0$$

$$\Rightarrow -(v^T y)^T (v^T y) \leq 0$$

$$-\text{"}(v^T y)^2\text{"} \leq 0$$

Midterm 2014 ✓ 1/3

1(a)

$$\theta = (X^T X)^{-1} X^T \vec{y} \quad \text{(from Notes - Normal Equation)}$$

$$\Rightarrow J(\theta) = (X\theta - \vec{y})(X\theta - \vec{y})$$

Side calc: $(X\theta - \vec{y}) = (X \overbrace{(X^T X)^{-1}}^{= I \text{ according to orthonormality assumption}} X^T \vec{y} - \vec{y})$

$$= (X X^T \vec{y} - \vec{y})$$

$$\Rightarrow \boxed{J(\theta) = (X X^T \vec{y} - \vec{y})^T (X X^T \vec{y} - \vec{y})}$$

b)

$$\nabla_\theta \tilde{J}(\theta) = \nabla_\theta J(\theta) + \theta_p.$$

$$\theta_{new} = \tilde{X}^T \vec{y}$$

$$\Rightarrow \tilde{J}(\theta_{new}) = (\tilde{X} \tilde{X}^T \vec{y} - \vec{y})^2$$

$$\Rightarrow \Delta = (\tilde{X} \tilde{X}^T \vec{y} - \vec{y}) - (X X^T \vec{y} - \vec{y})$$

$$= \tilde{X} \tilde{X}^T \vec{y} - X X^T \vec{y}$$

PE 2014 2/3

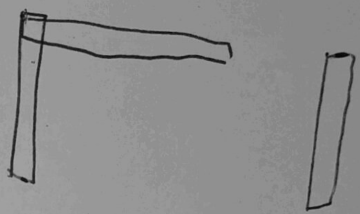
1b)

$\cancel{X\theta}$

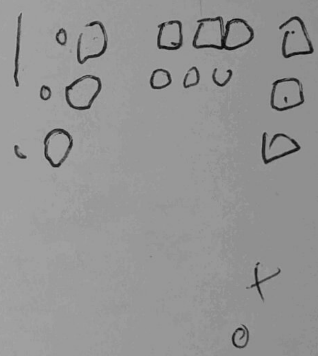
$\cancel{X\theta = \tilde{X}\theta_{new} -}$

$$\tilde{X}\theta_{new} = X\theta + \vec{v}\cdot\rho$$

$$\vec{v}\rho = \tilde{X}\theta_{new} - X\theta$$



$$\tilde{X}\tilde{X}^T = XX^T + vv^T$$



$$\Rightarrow \tilde{J}(\theta_{new}) =$$

$$= \left(\left(XX^T + vv^T\right)\vec{y} - \vec{y}\right)^2$$

$$= \left(XX^T\vec{y} + vv^T\vec{y} - \vec{y}\right)^2$$

$$\left(XX^T\vec{y} - \vec{y}\right)^T\left(XX^T\vec{y} - y\right) = \left(XX^T\vec{y}\right)^T\left(XX^T\vec{y}\right) - 2\left(XX^T\vec{y}\right)y + y^Ty$$