Math 189 Hw 2
Joseph Gardi
Feb 11, 2019

1) a)

$$\frac{d\sigma(x)}{dx} = -(1+e^{-x})^{-2}(-e^{-x})$$

$$= \frac{e^{-x}}{(1+e^{-x})^2}$$

$$= \frac{1+e^{-x}-1}{(1+e^{-x})^2} = \frac{1+e^{-x}}{(1+e^{-x})^2} - \frac{1}{(1+e^{-x})^2}$$

$$= \frac{1}{1+e^{-x}} - \frac{1}{(1+e^{-x})^2}$$

$$= \frac{1}{1+e^{-x}}\left(1 - \frac{1}{1+e^{-x}}\right) = \sigma(x)[1-\sigma(x)]$$

b) ~~Let y_i be~~ Let $y_i \in \{-1, 1\}$,
Log likelihood is,

$$-\sum_i \log(1+\exp(-y_i w^T x_i))$$

~~The gradient with respect to w_i is~~
The derivative with respect to w_j

$$-\sum_i \frac{1}{1+\exp(-y_i w^T x_i)}(-y_i w_j x_{i,j}) = \boxed{\sum_i \frac{y_i w_j x_{ij}}{1+\exp(-y_i w^T x_i)}}$$

~~I can find the derivative with the chain rule,~~

$\cancel{\frac{\partial}{\partial w_j} \sum_i \log y_i}$

c) This happens to be the spectral decomposition of H! What a coincidence. That means S contains the eigenvalues and H is positive semidefinite when all ~~ei~~ eigenvalues are ≥ 0.

see back

1.c continued) so eigenvalues of H are

$$\mu_i(1-\mu_i) = \mu_i - \mu_i^2$$

$$\geq 0 \quad (\text{since } \mu_i \in [0,1])$$

so $H \succeq 0$

2) I used the solution for this

It must integrat to Z so

$$\int_{-\infty}^{\infty} \frac{1}{Z} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = \frac{1}{Z} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = 1$$

$$\text{so } Z = \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx$$

Now consider

$$Z^2 = \int_{\mathbb{R}} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \int_{\mathbb{R}} \exp\left(-\frac{y^2}{2\sigma^2}\right) dy$$

$$= \iint_{\mathbb{R}^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) dx\, dy$$

$$= \int_0^{\infty} \int_0^{2\pi} \exp\left(-\frac{r^2}{2\sigma^2}\right) r\, d\theta\, dr$$

$$= 2\pi \int_0^{\infty} \exp\left(\frac{-r^2}{2\sigma^2}\right) r\, dr$$

$$= -2\pi\sigma^2 \int_0^{\infty} \exp\left(-\frac{r^2}{2\sigma^2}\right)\left(-\frac{r}{\sigma^2}\right) dr$$

$$= -2\pi\sigma \exp\left(-\frac{r^2}{2\sigma^2}\right)\Big|_0^{\infty} = -2\pi\sigma^2(-1) = 2\pi\sigma^2$$

inued)

So $\boxed{Z = \sqrt{2\pi\sigma^2}}$

3) a) $\underset{w}{\operatorname{argmax}} \sum^N \log \mathcal{N}(y_i \mid w_0 + w^T x_i, \sigma^2) + \sum^D \log \mathcal{N}(w_j \mid 0, \tau^2)$

$= \underset{w}{\operatorname{argmax}} \sum^N \log \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - (w_0 + w^T x_i))^2}{2\sigma^2}\right) + \sum^D \log \frac{1}{\sqrt{2\pi\tau^2}} \exp\left(-\frac{w_j^2}{2\tau^2}\right)$

$= \underset{w}{\operatorname{argmax}}\ N \log \frac{1}{\sqrt{2\pi\sigma^2}} + \sum^N \left(-\frac{(y_i - (w_0 + w^T x_i))^2}{2\sigma^2}\right) + D \log \frac{1}{\sqrt{2\pi\tau^2}} + \sum^D \left(-\frac{w_j^2}{2\tau^2}\right)$

$= \underset{w}{\operatorname{argmax}} -\sum^N (y_i - (w_0 + w^T x_i))^2 - \lambda \sum^D w_j^2$ (multiplied by $2\sigma^2$ and drop constants)

$= \underset{w}{\operatorname{argmin}} \sum^N (y_i - (w_0 + w^T x_i))^2 + \lambda \|w\|_2^2$

b) First take the derivative. ~~and set to 0~~ ~~then set to 0 and solve~~.

$2A^T(Ax - b) + 2\tau^2 x = (A^T A + \tau^2 I)x - A^T b$

Now set to 0 and solve

$(A^T A + \tau^2 I)x - A^T b = 0$

$\Rightarrow \boxed{x = (A^T A + \tau^2 I)^{-1} A^T b}$

Gradient with respect to X is

$$2A^T(AX - b\vec{1} - y) + 2\tau^2 X$$

Gradient with respect to b is

$$2\vec{1}^T(AX + b\vec{1} - y)$$

I had to use the solution for this.

d) I used the solution for this,

$$\min (Ax + b\vec{1} - y)^T (Ax + b\vec{1} - y) + (Tx)^T (Tx)$$

$$= \min (x^T A^T + b\vec{1}^T - y^T)(Ax + b\vec{1} - y) + x^T T^T T x$$

$$= \min x^T A^T A x + 2b\vec{1}^T A x - 2y^T A x - 2b\vec{1}^T y + b^2 n + y^T y + x^T T^T T x$$

Gradients are,

$$\nabla_x f = 2A^T A x + 2bA^T\vec{1} - 2A^T y + 2T^T T x$$

$$\nabla_b f = 2\vec{1}^T A x - 2\vec{1}^T y + 2bn$$

Set $\nabla_b f = 0$ and solve

$$2\vec{1}^T Ax - 2\vec{1}^T y + 2bn = 0$$

$$\Rightarrow b = \frac{2\vec{1}^T y - 2\vec{1}^T Ax}{2n} = \frac{\vec{1}^T (y - Ax)}{n}$$

Set $\nabla_x f$ and solve

$$2A^T A x + 2\frac{\vec{1}^T (y - Ax)}{n} A^T \vec{1} - 2A^T y + 2T^T T x = 0$$

$$\cancel{2A^T A x + \frac{2}{n}\vec{1}^T y A^T \vec{1} - \frac{2}{n}\vec{1}^T A x A^T \vec{1} - 2A^T y + 2T^T T x = 0}$$

$$(A^T A + T^T T)x + \frac{\vec{1}^T (y - Ax)}{n} A^T \vec{1} - A^T y = 0$$

$$(A^T A + T^T T)x + \frac{1}{n} A^T \vec{1}\vec{1}^T y - \frac{1}{n} A^T \vec{1}\vec{1}^T A x - A^T y = 0$$

$$\Rightarrow \left[A^T A + T^T T - \frac{1}{n} A^T \vec{1}\vec{1}^T A\right] x = A^T y - \frac{1}{n} A^T \vec{1}\vec{1}^T y$$

$$\Rightarrow x = \left[A^T \left(I - \frac{1}{n}\vec{1}\vec{1}^T\right) A + T^T T\right]^{-1} A^T \left(I - \frac{1}{n}\vec{1}\vec{1}^T\right) y$$