# Problem 1

Formulation :

$$p = x_1 \exp\left(A_{12}\left(\frac{A_{21}x_2}{A_{12}x_1 + A_{21}x_2}\right)^2\right)p_1^{sat} + x_2 \exp\left(A_{21}\left(\frac{A_{12}x_1}{A_{12}x_1 + A_{21}x_2}\right)^2\right)p_2^{sat}$$

The goal is to fit $A_{12}$ and $A_{21}$ with least squares

$$\min_{A_{12}, A_{21}} \sum_{i=1}^{N} \left(y_i - p(x_i, A_{12}, A_{21})\right)^2$$

subject to $x_1 + x_2 = 1$

$$\log_{10}(p_{sat}) = a_1 - \frac{a_2}{T + a_3}, \quad a_1, a_2, a_3 \text{ in Table}$$

$$T = 20$$

All of the constraints can be solved and substituted into objective functions

## Saturation Pressures

$$p_1^{sat} = p_{water}^{sat} = 10^{\left(8.07131 - \frac{1730.63}{20 + 233.426}\right)} = 17.4733$$

$$p_2^{sat} = p_{1,4\ dioxane}^{sat} = 10^{\left(7.43155 - \frac{1554.679}{20 + 240.337}\right)} = 28.8241$$

## Binary

$$x_1 + x_2 = 1$$

$$x_2 = 1 - x_1$$

$$\min_{A_{12}, A_{21}} \sum_{i=1}^{N} \left( y_i - \left[ x_{1,i} \exp\left( A_{12} \left( \frac{A_{21} X_{2,i}}{A_{12} X_{1,i} + A_{21} X_{2,i}} \right)^2 \right) p_1^{sat} + X_{2,i} \exp\left( A_{21} \left( \frac{A_{12} X_{1,i}}{A_{12} X_{1,i} + A_{21} X_{2,i}} \right)^2 \right) p_2^{sat} \right] \right)^2$$

$$\boxed{\min_{A_{12}, A_{21}} \sum_{i=1}^{N} \left( y_i - \left[ X_{1,i} \exp\left( A_{12} \left( \frac{A_{21} (1-X_{1,i})}{A_{12} X_{1,i} + A_{21} (1-X_{1,i})} \right)^2 \right) p_{wat}^{sat} + (1-X_{1,i}) \exp\left( A_{21} \left( \frac{A_{12} X_{1,i}}{A_{12} X_{1,i} + A_{21} (1-X_{1,i})} \right)^2 \right) p_{diox}^{sat} \right] \right)^2}$$

This is a nonlinear problem

so we will try to

$$\min_{a} \frac{1}{2} f(a)^T f(a)$$

where

$$f(a) = \begin{bmatrix} y_1 - p(x_1, a) \\ \vdots \\ y_n - p(x_n, a) \end{bmatrix}$$

The gradient

$$\frac{\partial F}{\partial a} = J^T f(a)$$

$$\begin{bmatrix} \frac{\partial f_1}{\partial a_1} & \frac{\partial f_1}{\partial a_2} & \cdots & \frac{\partial f_1}{\partial a_p} \\ \vdots & & & \\ \frac{\partial f_n}{\partial a_1} & \cdots & \cdots & \frac{\partial f_n}{\partial a_p} \end{bmatrix} f(a)$$

$$\boxed{a_{k+1} = a_k - \left( J^T J + \lambda I \right)^{-1} J^T f(a)}$$

# Problem 3

$$p(y_i;\, x, \theta) = \frac{1}{1+\exp(-y\theta^T x)}$$

Assuming $X$ is iid, the liklihood is:

$$L(\theta, D) = \prod_{i=1}^{N} \frac{1}{1+\exp(-y\theta^T x)} = \frac{1}{\prod_{i=1}^{N} \left[1+\exp(-y\theta^T x)\right]}$$

where

$$D = \{(x_i, y_i)\}_{i=1}^{N}$$

Taking the negative log, to get the negative log liklihood function gives:

$$\ell(\theta, D) = -\left(\cancelto{0}{\ln(1)} - \ln\left[\prod_{i=1}^{N}\left[1+\exp(-y_i\theta^T x_i)\right]\right]\right)$$

$$= \sum_{i=1}^{N} \ln\left(1+\exp(-y_i\theta^T x_i)\right)$$

To find the max likelihood of $\theta$, we need to minimize $\ell(\theta, D)$, so take derivative of $\ell$ w.r.t. $\theta$

$$\frac{\partial \ell}{\partial \theta} = \sum_{i=1}^{N} \frac{1}{1+\exp(-y_i\theta^T x_i)}\left(-y_i x_i\left(\exp\left(-y_i\,\theta^T x_i\right)\right)\right)$$

$$= \sum_{i=1}^{N} \frac{-y_i x_i \left( \exp(-y_i \theta^T x_i) \right)}{1 + \exp(-y_i \theta^T x_i)} = 0$$

variable substitute

$$r = \exp(-y_i \theta^T x_i), \quad -y_i x_i = a$$

$$= \sum_{i=1}^{N} a \left( \frac{r}{1+r} \right) = \frac{\partial \ell}{\partial \theta}$$

$$\frac{\partial^2 \ell}{\partial \theta^2} = \sum_{i=1}^{N} a \frac{\partial}{\partial r}\left( \frac{r}{1+r} \right) \cdot \frac{\partial r}{\partial \theta}$$

$$\frac{\partial}{\partial r}\left( \frac{r}{1+r} \right) = \frac{1}{(1+r)^2}$$

$$\frac{\partial r}{\partial \theta} = -y_i x_i \exp(-y_i \theta^T x_i)$$

$$= \sum_{i=1}^{N} \frac{(y_i x_i)^2 \exp(-y_i \theta^T x_i)}{\left( 1 + \exp(-y_i \theta^T x_i) \right)^2}$$

$y_i$ is either $-1$ or $1$, so $y_i^2 = 1$

$$\frac{\partial^2 \ell}{\partial \theta^2} = \sum_{i=1}^{N} \frac{(x_i)^2 \exp(-y_i \theta^T x_i)}{\left( 1 + \exp(-y_i \theta^T x_i) \right)^2}$$

$x_i^2$ is always positive, the exp function is always positive, therefore

$$\frac{\partial^2 \ell}{\partial \theta^2} > 0$$

Because the Hessian is positive definite, then $\ell$ is strictly convex, so there is only one $\theta_{max}$ and it is unique.