1. Answer

(1) $q(\theta_i|\psi_i) = \frac{1}{Z}\exp\{E_{q_{j\neq i}}[\ln p(X,\theta_1,\cdots,\theta_m)]$

$$q(w,\lambda,\alpha_1,\cdots,\alpha_d) = q(w)q(\lambda)\prod_{k=1}^{d} q(\alpha_k).$$

λ: $q(\lambda) \propto \exp\{E_{q(w)}[\ln p(y|w,\lambda,x) + \ln p(w|\alpha_1,\cdots,\alpha_d)] + \ln p(\lambda)\}$

$$\propto \exp\{E_{q(w)}[\sum_{i=1}^{N}\ln p(y_i|x_i,\lambda,w)] + \ln p(\lambda)\}$$

$$\propto \prod_{i=1}^{N}\lambda^{\frac{1}{2}}e^{-\frac{\lambda}{2}E_{q(w)}[(y_i-x_i^Tw)^2]}\lambda^{-1}e^{-f\lambda}$$

We assume ~~Ett~~ $q(\lambda) \sim Gamma(\lambda|e',f')$

$$\begin{cases} e' = e_0 + \frac{N}{2} \\ f' = f_0 + \frac{1}{2}\sum_{i=1}^{N}E_{q(w)}[(y_i-x_i^Tw)^2] \end{cases}$$

$$E_{q(w)}[(y_i-x_i^Tw)^2] = (y_i-x_i^T\mu')^2 + x_i^T\Sigma'x_i$$

w: $q(w) \propto \exp\{E_{q(\lambda)}[\ln(y|x,\lambda,w) + \ln p(\lambda)] + E_{q(\sum_{i=1}^{d}\alpha_i)}[\sum_{i=1}^{d}\ln p(\alpha_k) + \ln p(w|\alpha_1,\cdots,\alpha_d)]\}$

$$\propto \exp\{\sum_{i=1}^{N}E_{q(\lambda)}\ln p(y_i|x_i,\lambda,w) + E_{q(\sum_{i=1}^{d}\alpha_i)}[\ln p(w|\alpha_1,\cdots,\alpha_d)]\}$$

$$\propto [\prod_{i=1}^{N}e^{\frac{1}{2}E(\ln\lambda)}e^{-E_{q(\lambda)}[\lambda]/2\cdot(y_i-x_i^Tw)}]\prod_{k=1}^{d}e^{-\frac{E_q(\alpha_i)}{2}w_i^2}$$

We assume $q(w) \sim Normal(w|\mu',\Sigma')$

$$\begin{cases} \Sigma' = [diag(E_q(\alpha_1),\cdots,E_q(\alpha_d)) + E_{q(\lambda)}[\lambda]\cdot\sum_{i=1}^{N}x_ix_i^T]^{-1} \\ \mu' = \Sigma'(E_{q(\lambda)}[\lambda]\sum_{i=1}^{N}y_ix_i) \end{cases}$$

$$E_{q(\lambda)}[\lambda] = \frac{e'}{f'}$$

α_i: $q(\alpha_i) \propto \exp\{E_{q(w,\sum_{j\neq i}\alpha_j)}[\ln p(w|\alpha_1,\cdots,\alpha_d)] + \ln p(\alpha_i)\}$

$$\propto \alpha_i^{a_i-1}e^{-b_i\alpha_i}\cdot\exp\{\int q(w)\,\cancel{q}\,\ln\alpha_i^{\frac{1}{2}}(-\frac{\alpha_i}{2}w_i^2)\,dw$$

$$\propto \alpha_i^{a_i-\frac{1}{2}}e^{-b_i\alpha_i}\cdot\cancel{Z}\,e^{E(w_i^2)}$$

We assume $q(\alpha_i) \propto Gamma(\alpha_i|a_i',b_i')$

$$\begin{cases} a_i' = a_0 + \frac{1}{2} \\ b_i' = b_0 + \frac{1}{2}E(w_i^2) \end{cases}$$

$$E(w_i^2) = \Sigma'(i,i) + (\mu_i')^2$$

(2).

Inputs: Data and definitions $q(\lambda) = \text{Gamma}(\lambda | e', f')$, $q(w) = \text{Normal}(w | \mu', \Sigma')$

$q(\lambda_i) = \text{Gamma}(\lambda_i | a_i', b_i')$ $(i = 1, 2, \cdots, d)$.

Output: Values for $e', f', \mu', \Sigma', a_i', b_i'$ for $i = 1, 2, \cdots d$.

1. Initialize $e_0', f_0', \mu_0', \Sigma_0', a_{i0}', b_{i0}'$ for $i = 1, 2, \cdots, d$

2. For iteration $t = 1, 2, \cdots, T$

(1) Update $q(\lambda)$ by setting:

$$e_t' = e_0 + \frac{N}{2}$$

$$f_t' = f_0 + \frac{1}{2} \sum_{i=1}^{N} [(y_i - x_i^T \mu_{t-1}')^2 + x_i^T \Sigma_{t-1}' x_i]$$

(2) Update $q(w)$ by setting:

$$\Sigma_t' = [\text{diag}(E_{q(\lambda_1)}, \cdots, E_{q(\lambda_d)}) + E_{q(\lambda)}(\lambda) \cdot \sum_{i=1}^{N} x_i x_i^T]^{-1} \qquad (E_{q(\lambda)}(\lambda) = \frac{e_{t-1}'}{f_{t-1}'})$$

$$\mu_t' = \Sigma_{t-1}' (E_{q(\lambda)}(\lambda) \sum_{i=1}^{N} y_i x_i).$$

(3) Update $q(\lambda_i)$ for $i = 1, \cdots, d$ by setting:

$$a_{it}' = a_0 + \frac{1}{2}$$

$$b_{it}' = b_0 + \frac{1}{2} E(w_i^2) \qquad (E(w_i^2) = \Sigma_{t-1}'(i, i) + (\mu_{t-1}'(i))^2)$$

(4). Evaluate $L(e_t', f_t', \mu_t', \Sigma_t', a_{it}', b_{it}'$ for $i = 1, 2, \cdots d))$

to access convergence.

(3). $L(Q_i) = \int q(w, \alpha_1, \alpha_2, \cdots, \alpha_d) \ln \frac{p(y, w, \lambda | X)}{q(w, \alpha_1, \cdots \alpha_d | \lambda)}$ $\quad (Q_i = w, \lambda, \alpha_1, \quad , \alpha_d)$

$q(w, \alpha_1, \cdots, \alpha_d, \lambda) = q(w)\, q(\lambda) \prod_{i=1}^{d} \alpha_i$

$\cancel{P(w, \alpha_1, \quad, \alpha_d}$ $P(y, w, \lambda | X) = \prod_{i=1}^{N} P(y_i, w, \lambda | x_i)$

$= \prod_{i=1}^{N} P(y_i | w, \lambda, x_i) \cdot P(\lambda) \cdot P(w | \alpha_1, \cdots, \alpha_d) \cdot \prod_{k=1}^{d} P(\alpha_k)$

$$L(Q_i) = \sum_{i=1}^{N} \iint q(w) q(\lambda) \ln P(y_i | w, \lambda, x_i) + \int q(\lambda) \ln p(\lambda) + \sum_{i=1}^{d} q(\alpha_i) \ln P(\alpha_i)$$

$$+ \int q(w) q(\alpha_1) \cdots q(\alpha_d) \ln P(w | \alpha_1, \cdots, \alpha_d) - \int q(\lambda) \ln q(\lambda) - \int q(w) \ln q(w) - \sum_{i=1}^{d} \int q(\alpha_i) \ln q(\alpha_i)$$

$$= \frac{N}{2} (\psi(e') - \ln f') - \sum_{i=1}^{N} \frac{1}{2} \frac{e'}{f'} [(y_i - x_i^T) \mu']^2 + x_i^T \Sigma' x_i] + const$$

$$+ (e_0 - 1)(\psi(e') - \ln f') - f_0 \cdot \frac{e'}{f'} + const$$

$$+ \{ -\sum_{i=1}^{d} \frac{E(\alpha_i)}{2} (\mu_i^2 + \Sigma'(i,i))] + \sum_{i=1}^{d} \frac{1}{2} \{ \psi(a_i') - \ln b_i'] + const$$

$$+ \sum_{i=1}^{d} \{ (a_0 - 1) \{ \psi(a_i') - \ln b_i') - b_0 \cdot \frac{a_i'}{b_i'}] + const$$

$$+ e' - \ln f' + \ln \Gamma(e') + (1 - e') \psi(e') + const$$

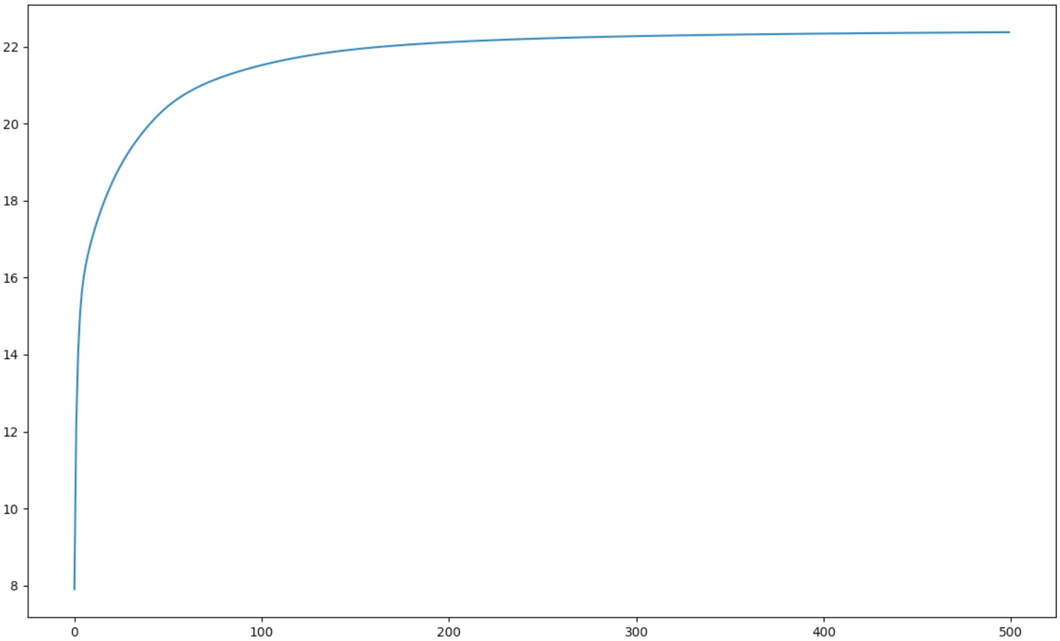
$$+ \frac{1}{2} \ln |\Sigma'| + const$$

$$+ \sum_{i=1}^{d} \{ a_i' - \ln b_i' + \ln \Gamma(a_i') + (1 - a_i') \psi(a_i')] + const$$

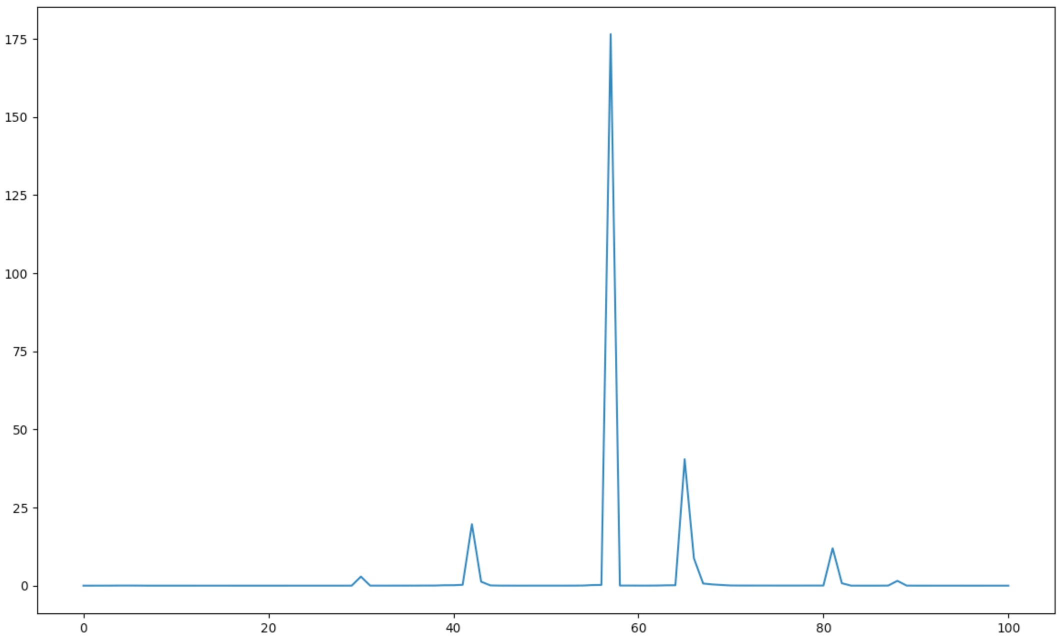
2.

Dataset1:

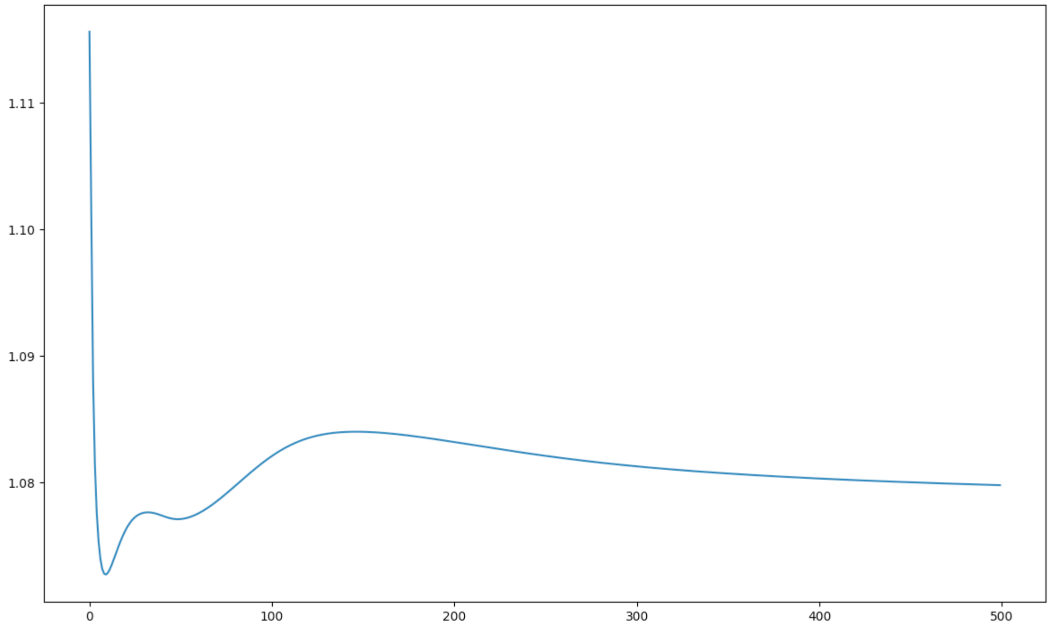
VI:



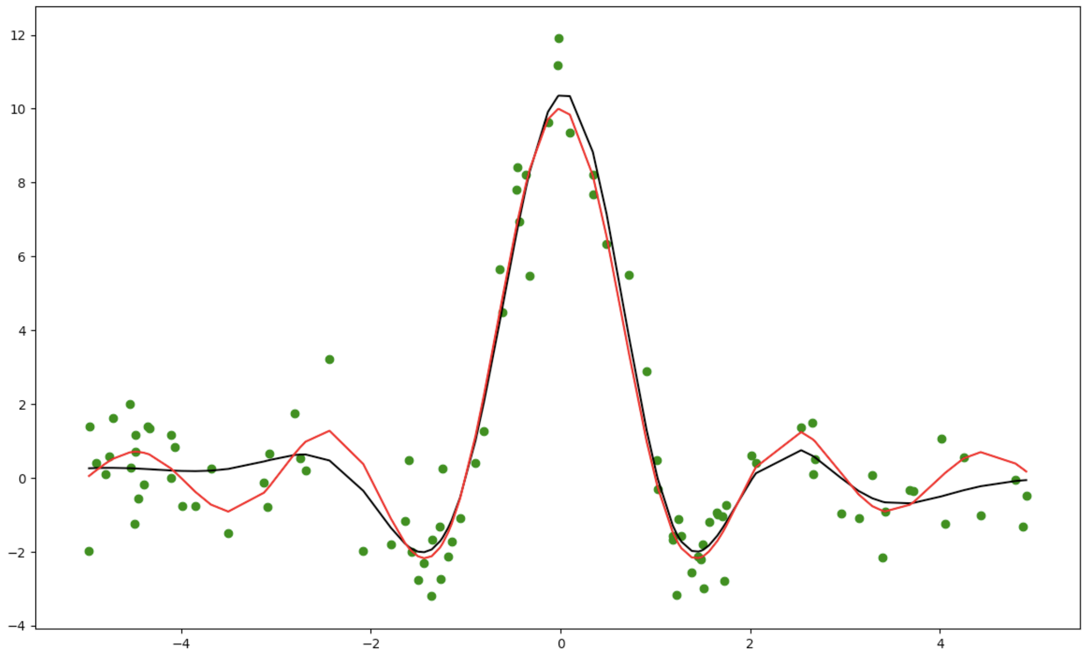
1/Eq_alpha:



1/Eq_lambda:



Omega:

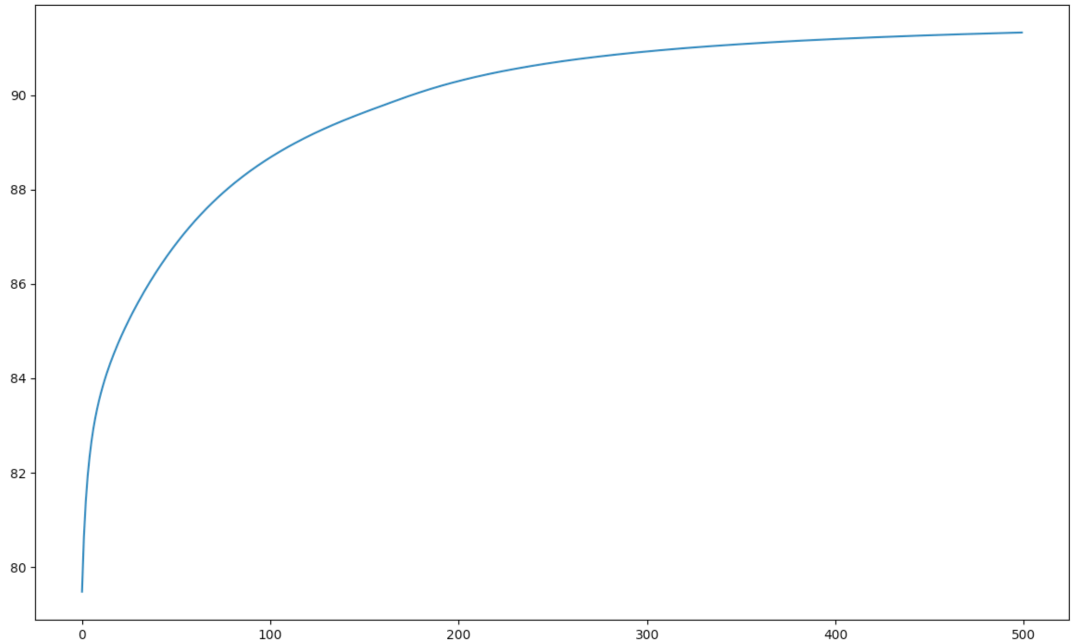


Red is plot(z_i; 10 * sinc(z_i)),   Black is plot(yi_predict, zi)

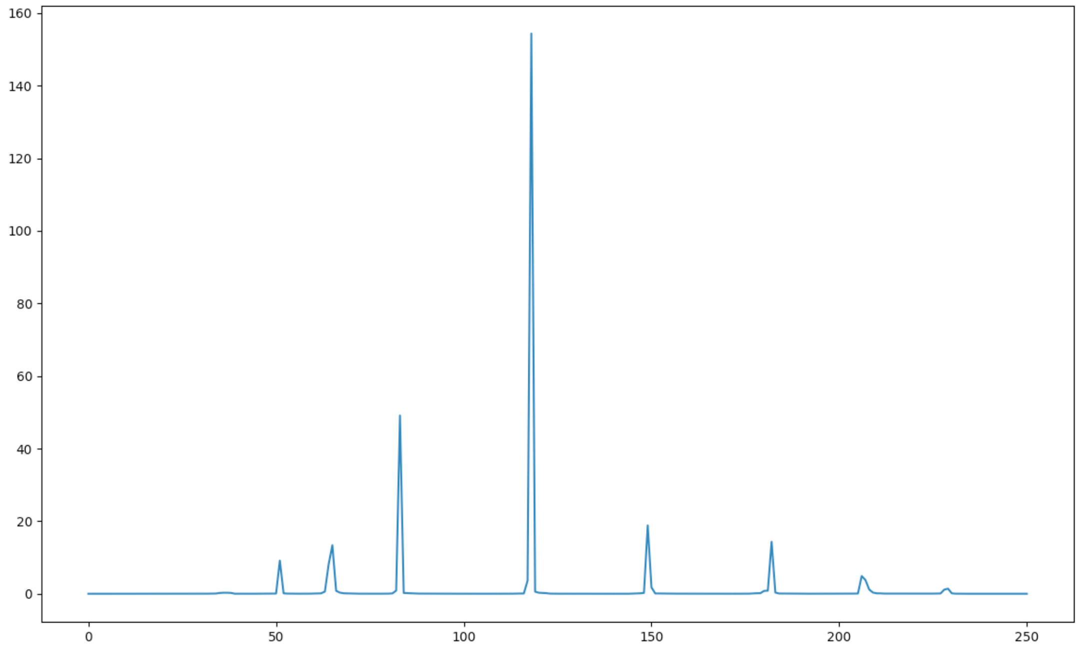
Green is scatter(yi, zi)

Dataset2:

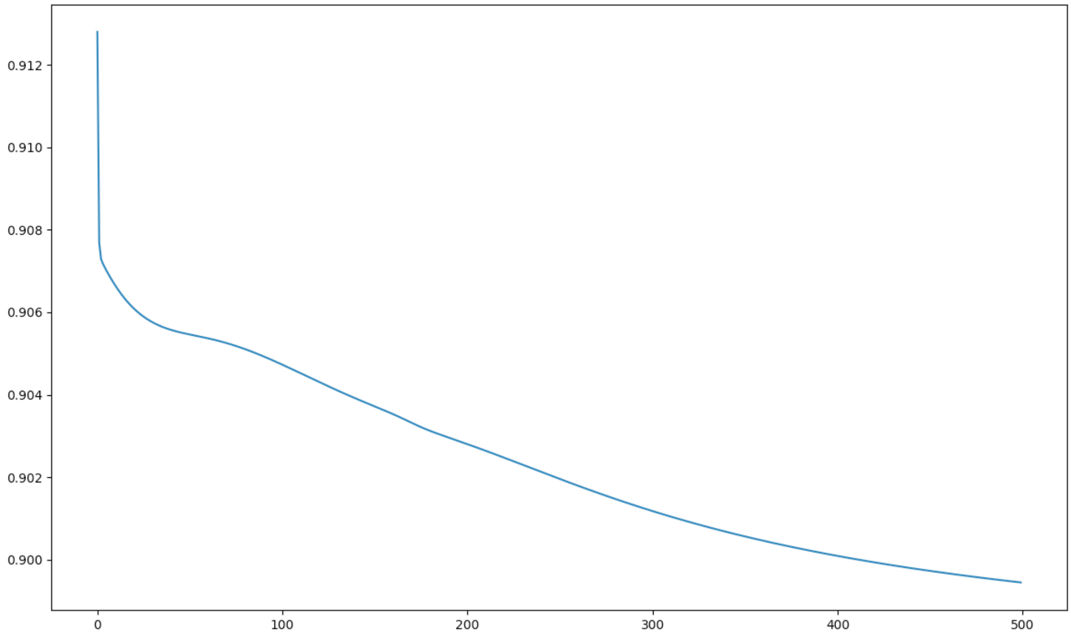
VI:



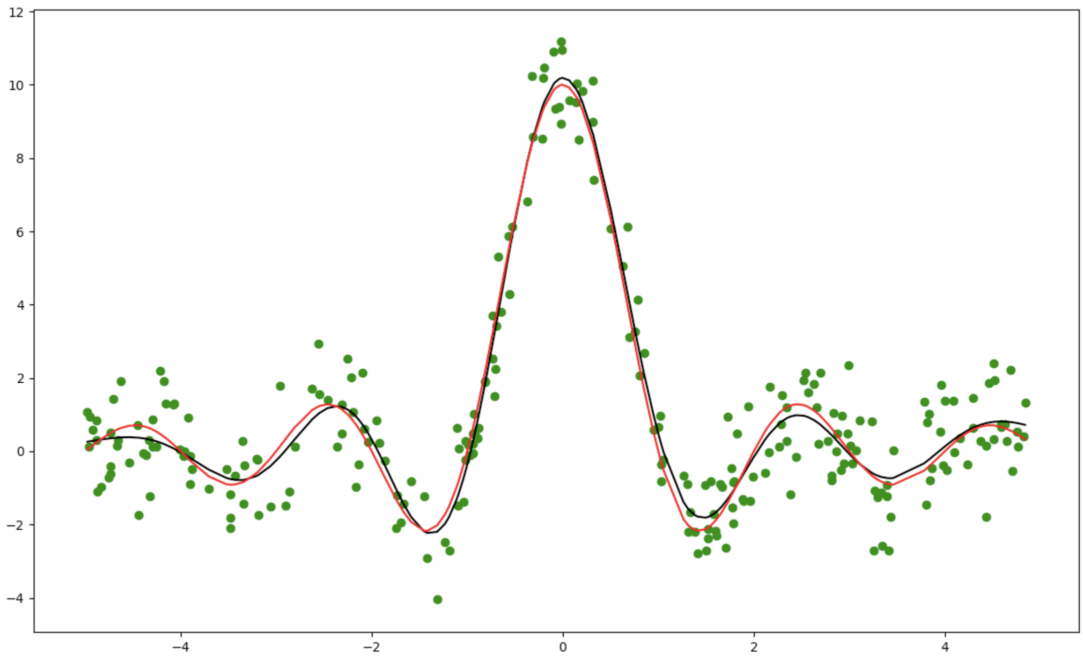
1/Eq_alpha:



1/Eq_lambda:

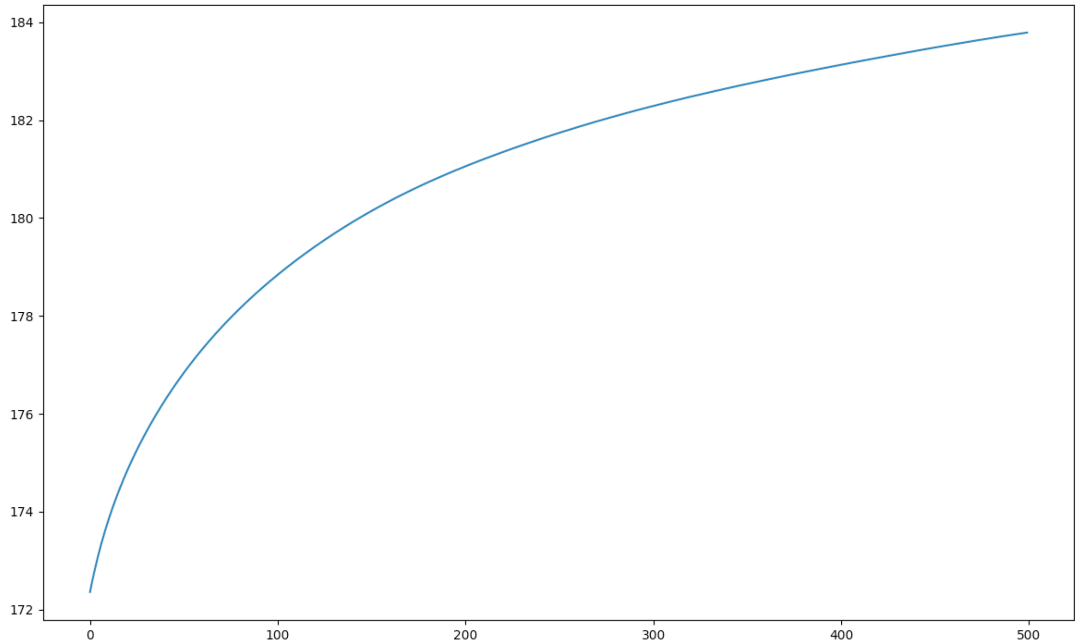


Omega:

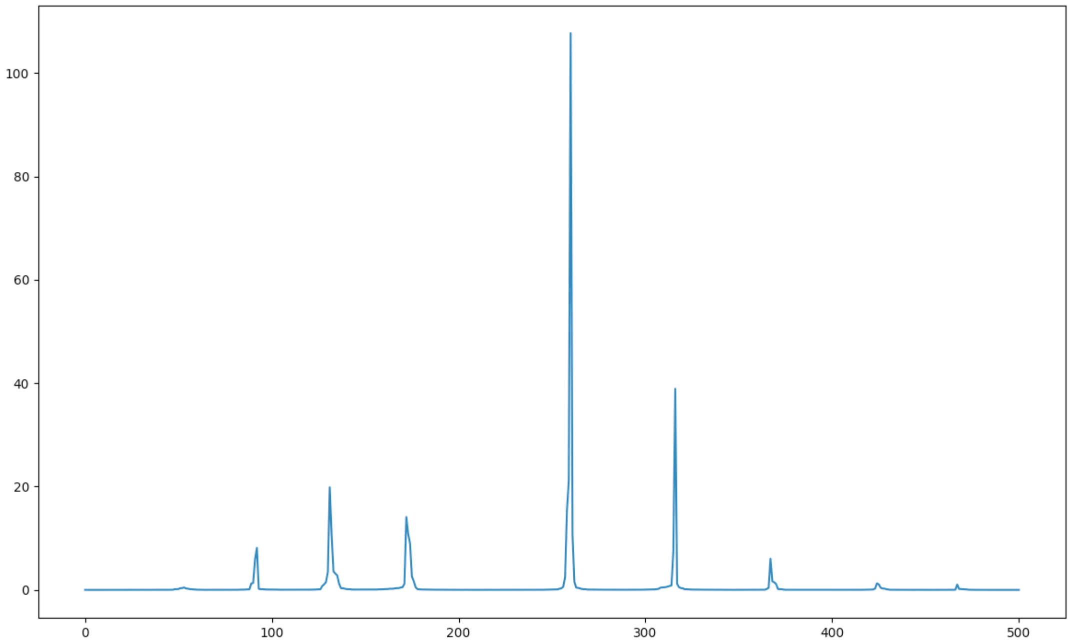


Dataset3:

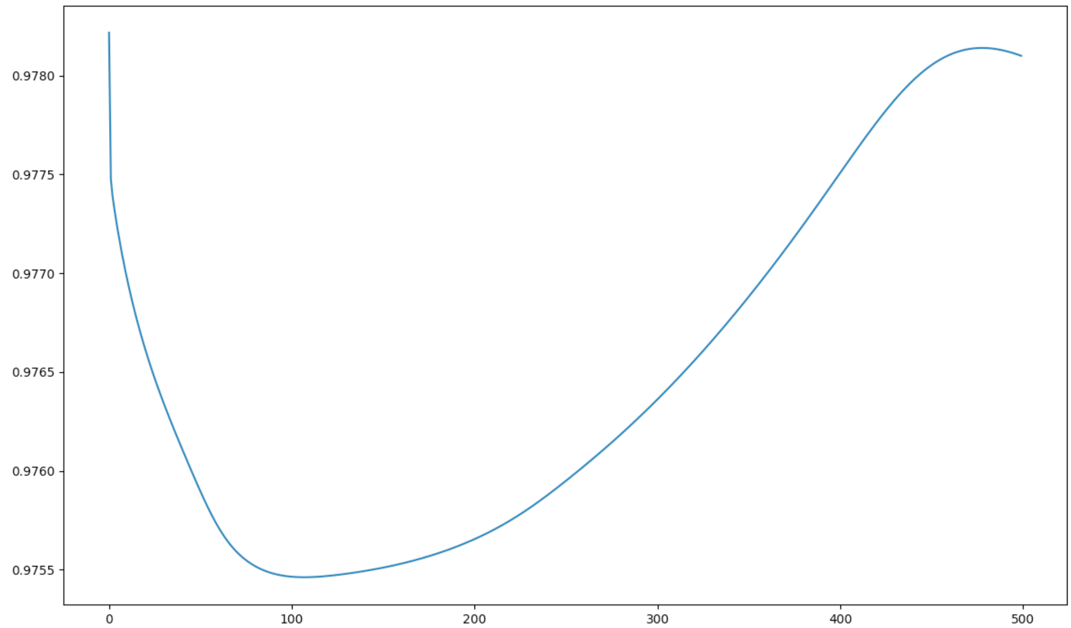
VI:



1/Eq_alpha:



1/Eq_lambda:



Omega: