Problem Set 1, Q3:

(1). $\langle \delta m \, \delta m^T \rangle = (A^T N^{-1} A)^{-1}$

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1/\sigma^2 & & & \\ & 1/\sigma^2 & & \\ & & \ddots & \\ & & & 1/\sigma^2 \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$\rightarrow \langle \delta m \, \delta m^T \rangle = \left(\frac{N}{\sigma^2}\right)^{-1} = \frac{\sigma^2}{N} \rightarrow \sqrt{\langle \delta m \, \delta m^T \rangle} = \frac{\sigma}{\sqrt{N}}$$

(2.) Now, if we get the error wrong on one half of data we have:

$$\langle \delta m \, \delta m^T \rangle_{new} = \left(\frac{N}{2} \frac{1}{\sigma^2} + \frac{N}{2} \left(\frac{2}{\sigma^2}\right) \right)^{-1} = \frac{2}{3} \frac{\sigma^2}{N}$$

This can also be $\frac{1}{2\sigma^2}$. It is not clear from the question that which one is the intention. Thus, I will stick to $\frac{2}{\sigma^2}$

$$\sqrt{\langle \delta m \, \delta m^T \rangle_{new}} = \frac{\sigma}{\sqrt{\frac{3}{2} N}} \rightarrow \frac{\sqrt{\langle m m^T \rangle_{new}}}{\sqrt{\langle m m^T \rangle}} = \sqrt{\frac{2}{3}}$$

(3.) If we assume that there is a weight associated with the data, the χ^2 will have the following form:

$$\chi^2 = (Wd - Am)^T N^{-1} (Wd - Am)$$

$$\nabla \chi^2 = -2A^TN^{-1}(Wd - Am) = 0 \rightarrow$$

$$A^TN^{-1}Am = A^TN^{-1}Wd$$

Now, if we set $d = d_t + n$, where n is our noise, we will get:

$$A^TN^{-1}Am_t = A^TN^{-1}Wd_t$$

Substracting the two, we get:

$$A^TN^{-1}A(m - m_t) = A^TN^{-1}W(d - d_t)$$

$$m - m_t = (A^TN^{-1}A)^{-1}A^TN^{-1}Wn$$

$$\langle \delta m \delta m^T \rangle = (A^TN^{-1}A)^{-1}A^TN^{-1}Wn\left[(A^TN^{-1}A)^{-1}A^TN^{-1}Wn\right]^T$$

$$\langle \delta m \delta m^T \rangle = (A^TN^{-1}A)^{-1}A^TN^{-1}W\underbrace{nn^T}_{N}W^TN^{-1}A(A^TN^{-1}A)^{-1}$$

$$\langle \delta m \delta m^T \rangle = \underbrace{(A^TN^{-1}A)^{-1}}A^TN^{-1}WNW^TN^{-1}A\underbrace{(A^TN^{-1}A)^{-1}}$$

Let's look at each term individually:

$$A^TN^{-1}A = [1 \;\cdots\; 1]_n \begin{bmatrix} 1/\sigma^2 & & & \\ & 1/\sigma^2 & & \\ & & \ddots & \\ & & & 1/\sigma^2 \end{bmatrix}_{n\times n} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_n$$

$$\rightarrow A^TN^{-1}A = \frac{N}{\sigma^2} \rightarrow (A^TN^{-1}A)^{-1} = \frac{\sigma^2}{N}$$

$$WNW^T = \begin{bmatrix} w_1 & & \\ & w_2 & \\ & & \ddots \end{bmatrix} \begin{bmatrix} \sigma^2 & & \\ & \ddots & \\ & & \end{bmatrix} \begin{bmatrix} w_1 & & \\ & \ddots & \\ & & \end{bmatrix}$$

$$\begin{bmatrix} \ddots & \\ & w_n \end{bmatrix}_{n\times n} \begin{bmatrix} \ddots & \\ & \sigma^2 \end{bmatrix}_{n\times n} \begin{bmatrix} \ddots & \\ & w_n \end{bmatrix}_{n\times n}$$

$$\rightarrow WNW^T = \begin{bmatrix} w_1^2\sigma^2 & & \\ & \ddots & \\ & & w_n^2\sigma^2 \end{bmatrix}_{n\times n}$$

$$\rightarrow N^{-1}WN\;W^TN^{-1} = \begin{bmatrix} w_1^2/\sigma^2 & & \\ & \ddots & \\ & & w_n^2/\sigma^2 \end{bmatrix}_{n\times n}$$

$$\rightarrow A^TN^{-1}WN\;W^TN^{-1}A = [1 \cdots 1]_n \begin{bmatrix} w_1^2/\sigma^2 & & \\ & \ddots & \\ & & w_n^2/\sigma^2 \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_n$$

$$\rightarrow A^TN^{-1}WN\;W^TN^{-1}A = \frac{N}{\sigma^2}\sum_{i=1}^{n} w_i^2$$

Now that we have the value for different compartments, we can combine them:

$$\langle \delta_m \delta_m^T \rangle = \frac{\sigma^2}{N} \times \frac{N}{\sigma^2} \times \sum_{i=1}^{N} w_i^2 \times \frac{\sigma^2}{N} = \frac{\sigma^2}{N}\sum_{i=1}^{N} w_i^2$$

$$\sqrt{\langle \delta_m \delta_m^T \rangle} = \sqrt{\frac{\sigma^2}{N}\sum w_i^2}$$

So in order to find the exact value, we need to evaluate $\sum w_i^2$.

We know that for 1% of the data, this factor is $1/100$. We are also aware that: $\sum w_i = 1$

$$\sum_{i=1}^{N} w_i = 1 \rightarrow 0.99N \underbrace{w} + 0.01N \times \underbrace{0.01w} = 1$$

$$0.99Nw + 10^{-4}Nw = 1 \rightarrow 0.9901\,Nw = 1 \rightarrow$$

$$Nw = 1.0099 \rightarrow w = \frac{1.0099}{N}$$

$$\rightarrow \sum_{i=1}^{N} w_i^2 = 0.99N \times \left(\frac{1.0099}{N}\right)^2 + 0.01N \times 10^{-4} \times \left(\frac{1.0099}{N}\right)^2$$

$$\rightarrow \sum_{i=1}^{N} w_i^2 = \frac{1.0096990299}{N} + \frac{0.0000010199}{N} = \frac{\overbrace{1.0097000498}^{\equiv A}}{N}$$

$$\sqrt{\langle \delta m \, \delta m^T \rangle} = \sqrt{\frac{\sigma^2}{N} \times \frac{A}{N}} = \boxed{1.0048383203 \, \frac{\sigma}{N}}$$

$$\frac{\sqrt{\langle \delta m \, \delta m^T \rangle \text{new}}}{\sqrt{\langle \delta m \, \delta m^T \rangle}} = \frac{\sqrt{A}\, \cancel{\sigma}/N}{\cancel{\sigma}/\sqrt{N}} = \sqrt{A}\, \frac{\sqrt{N}}{N} = \sqrt{\frac{A}{N}}$$

Thus, the value of difference is very small specically if N is larg.

4. By comparing these two cases, we can infer that mistakes who affect larger number of datas make greater concerns rather than those who have a weight $\gg 1$ or $\ll 1$ but affect a small group.