Statistical Analysis of Data (Addendum)

Experiment AA-P

To prove that

$$\langle \mathbf{a} \rangle = \boldsymbol{\alpha} \tag{1}$$

where

$$\mathbf{a} = [\mathbf{X}]^{-1}\mathbf{Y} \tag{2}$$

first note that in taking the expectation value only the y_i are relevant. Because $[\mathbf{X}]$ (or its inverse) is independent of the y_i, it can be factored out of the expectation value giving

$$\langle \mathbf{a} \rangle = [\mathbf{X}]^{-1}\langle \mathbf{Y} \rangle \tag{3}$$

The expectation value for the $\mathbf{Y}$ vector is

$$\langle \mathbf{Y} \rangle = \begin{pmatrix} \langle Y_1 \rangle \\ \langle Y_2 \rangle \end{pmatrix} \tag{4}$$

The defining equation for Y_m is

$$Y_m = \sum_{i=1}^{N} \frac{y_i f_m(x_i)}{\sigma_i^2} \tag{5}$$

Taking its expectation value gives

$$\langle Y_m \rangle = \left\langle \sum_{i=1}^{N} \frac{y_i f_m(x_i)}{\sigma_i^2} \right\rangle = \sum_{i=1}^{N} \frac{f_m(x_i)}{\sigma_i^2} \langle y_i \rangle \tag{6}$$

Recall that y_i is a sample from a Gaussian pdf of mean given by the fitting funnction $\mu_i = F(x_i)$ or

$$\mu_i = \alpha_1 f_1(x_i) + \alpha_2 f_2(x_i) \tag{7}$$

Substituting in Eq. 6 $\langle y_i \rangle = \mu_i$ in this form gives

$$\begin{aligned}
\langle Y_m \rangle &= \sum_{i=1}^{N} \frac{f_m(x_i)}{\sigma_i^2}(\alpha_1 f_1(x_i) + \alpha_2 f_2(x_i)) && (8) \\
&= \alpha_1 \sum_{i=1}^{N} \frac{f_m(x_i) f_1(x_i)}{\sigma_i^2} + \alpha_2 \sum_{i=1}^{N} \frac{f_m(x_i) f_2(x_i)}{\sigma_i^2} && (9) \\
&= \alpha_1 x_{1m} + \alpha_2 x_{2m} && (10)
\end{aligned}$$

which is simply the component form of the equation

$$\langle \mathbf{Y} \rangle = [\mathbf{X}]\boldsymbol{\alpha} \tag{11}$$

Using this in Eq. 3 gives

$$\begin{aligned} \langle \mathbf{a} \rangle &= [\mathbf{X}]^{-1}[\mathbf{X}]\boldsymbol{\alpha} & (12) \\ &= \boldsymbol{\alpha} & (13) \end{aligned}$$

thus completing the proof.

To prove that

$$[\mathbf{X}]^{-1} = [\boldsymbol{\sigma}_a^2] \tag{14}$$

it will be notationally convenient useful to define the row vectors $\mathbf{a}^T$, $\boldsymbol{\alpha}^T$, $\mathbf{Y}^T$

$$\begin{aligned} \mathbf{a}^T &= \begin{pmatrix} a_1 & a_2 \end{pmatrix} & (15) \\ \boldsymbol{\alpha}^T &= \begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix} & (16) \\ \mathbf{Y}^T &= \begin{pmatrix} Y_1 & Y_2 \end{pmatrix} & (17) \end{aligned}$$

The ordered product of a row vector and its corresponding column vector is then a scalar quantity (inner or dot product). For example,

$$\begin{aligned} \boldsymbol{\alpha}^T\boldsymbol{\alpha} &= \begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} & (18) \\ &= \alpha_1^2 + \alpha_2^2 & (19) \end{aligned}$$

while the ordered product of a column vector and its corresponding row vector is a square matrix (outer product). For example,

$$\begin{aligned} \boldsymbol{\alpha}\boldsymbol{\alpha}^T &= \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix} & (20) \\ &= \begin{pmatrix} \alpha_1^2 & \alpha_1\alpha_2 \\ \alpha_2\alpha_1 & \alpha_2^2 \end{pmatrix} & (21) \end{aligned}$$

The defining equation for $[\boldsymbol{\sigma}_a^2]$ is given explicitly by

$$\begin{aligned} [\boldsymbol{\sigma}_a^2] &= \begin{pmatrix} \sigma_1^2 & \sigma_{12}^2 \\ \sigma_{12}^2 & \sigma_1^2 \end{pmatrix} & (22) \\ &= \begin{pmatrix} \langle (a_1-\alpha_1)^2 \rangle & \langle (a_1-\alpha_1)(a_2-\alpha_2) \rangle \\ \langle (a_1-\alpha_1)(a_2-\alpha_2) \rangle & \langle (a_1-\alpha_1)^2 \rangle \end{pmatrix} & (23) \end{aligned}$$

Then, bringing the expectation value outside the matrix and expressing the matrix as an outer product of a column vector $\mathbf{a} - \boldsymbol{\alpha}$ and its corresponding row vector $\mathbf{a}^T - \boldsymbol{\alpha}^T$ gives

$$\begin{aligned} [\boldsymbol{\sigma}_a^2] &= \langle (\mathbf{a} - \boldsymbol{\alpha})(\mathbf{a}^T - \boldsymbol{\alpha}^T) \rangle & (24) \\ &= \langle \mathbf{a}\mathbf{a}^T \rangle - \langle \mathbf{a} \rangle \boldsymbol{\alpha}^T - \boldsymbol{\alpha}\langle \mathbf{a}^T \rangle + \boldsymbol{\alpha}\boldsymbol{\alpha}^T & (25) \\ &= \langle \mathbf{a}\mathbf{a}^T \rangle - \boldsymbol{\alpha}\boldsymbol{\alpha}^T & (26) \end{aligned}$$

To get from the second to third line we have used Eq. 1 $\langle \mathbf{a} \rangle = \boldsymbol{\alpha}$ and its equivalent for row vectors $\langle \mathbf{a}^T \rangle = \boldsymbol{\alpha}^T$. But we still need to calculate $\langle \mathbf{a}\mathbf{a}^T \rangle$

The transpose of the defining equation for $\mathbf{a}$ (Eq. 2) is

$$\mathbf{a}^T = \mathbf{Y}^T[\mathbf{X}]^{-1} \tag{27}$$

where use has been made of the fact that $[\mathbf{X}]$ and its inverse are symmetric and thus $[\mathbf{X}]^{-1}$ is equal to its transpose. Thus,

$$\mathbf{a}\mathbf{a}^T = [\mathbf{X}]^{-1}\mathbf{Y}\mathbf{Y}^T[\mathbf{X}]^{-1} \tag{28}$$

Factoring the $[\mathbf{X}]^{-1}$ matrices from the expectation value gives

$$\langle \mathbf{a}\mathbf{a}^T \rangle = [\mathbf{X}]^{-1}\langle \mathbf{Y}\mathbf{Y}^T \rangle[\mathbf{X}]^{-1} \tag{29}$$

Now, of course, the expectation value $\langle \mathbf{Y}\mathbf{Y}^T \rangle$ is needed. It has the explicit representation

$$\langle \mathbf{Y}\mathbf{Y}^T \rangle = \begin{pmatrix} \langle Y_1 Y_1 \rangle & \langle Y_1 Y_2 \rangle \\ \langle Y_1 Y_2 \rangle & \langle Y_2 Y_2 \rangle \end{pmatrix} \tag{30}$$

Using Eq. 5, the expectation value of the element $\langle Y_m Y_n \rangle$ becomes

$$\langle Y_m Y_n \rangle = \left\langle \left(\sum_{i=1}^{N} \frac{y_i f_m(x_i)}{\sigma_i^2} \right) \left(\sum_{i}^{N} \frac{y_i f_n(x_i)}{\sigma_i^2} \right) \right\rangle \tag{31}$$

Converting the product of the sums into two sums, one containing direct terms y_i^2 and one containing cross terms $y_i y_j$, $i \neq j$ gives

$$\langle Y_m Y_n \rangle = \left\langle \sum_{i=1}^{N} \frac{y_i^2 f_m(x_i) f_n(x_i)}{\sigma_i^4} + 2 \sum_{i \neq j}^{N} \frac{y_i y_j f_m(x_i) f_n(x_j)}{\sigma_i^2 \sigma_j^2} \right\rangle \tag{32}$$

The expectation value can now be taken term by term

$$\langle Y_m Y_n \rangle = \sum_{i=1}^{N} \frac{\langle y_i^2 \rangle f_m(x_i) f_n(x_i)}{\sigma_i^4} + 2 \sum_{i \neq j}^{N} \frac{\langle y_i y_j \rangle f_m(x_i) f_n(x_j)}{\sigma_i^2 \sigma_j^2} \tag{33}$$

The expectation value $\langle y_i^2 \rangle = \mu_i^2 + \sigma_i^2$ and, since the y_i are independent, the expectation value $\langle y_i y_j \rangle = \mu_i \mu_j$. Making these substitutions gives

$$\begin{aligned}
\langle Y_m Y_n \rangle &= \sum_{i=1}^{N} \frac{(\mu_i^2 + \sigma_i^2) f_m(x_i) f_n(x_i)}{\sigma_i^4} + 2 \sum_{i \neq j}^{N} \frac{\mu_i \mu_j f_m(x_i) f_n(x_j)}{\sigma_i^2 \sigma_j^2} && (34) \\
&= \sum_{i=1}^{N} \frac{f_m(x_i) f_n(x_i)}{\sigma_i^2} + \sum_{i=1}^{N} \frac{\mu_i^2 f_m(x_i) f_n(x_i)}{\sigma_i^4} + 2 \sum_{i \neq j}^{N} \frac{\mu_i \mu_j f_m(x_i) f_n(x_j)}{\sigma_i^2 \sigma_j^2} && (35) \\
&= \sum_{i=1}^{N} \frac{f_m(x_i) f_n(x_i)}{\sigma_i^2} + \left(\sum_{i=1}^{N} \frac{\mu_i f_m(x_i)}{\sigma_i^2} \right) \left(\sum_{i=1}^{N} \frac{\mu_i f_n(x_i)}{\sigma_i^2} \right) && (36)
\end{aligned}$$

where in the last line, the final two sums have been put back together as the product of two sums (the reverse of taking them apart as in going from Eq. 31 to Eq. 32 but now with μ_i's in place of the y_i's). The first term is the definition of x_{mn}, and when Eq. 7 is used to substitute for μ_i each of the sums in the product become

$$\begin{aligned}\sum_{i=1}^{N}\frac{\mu_i f_n(x_i)}{\sigma_i^2} &= \sum_{i=1}^{N}\frac{(\alpha_1 f_1(x_i)+\alpha_2 f_2(x_i))f_n(x_i)}{\sigma_i^2} && (37)\\ &= \alpha_1 x_{1n}+\alpha_2 x_{2n} && (38)\end{aligned}$$

Thus

$$\langle Y_m Y_n\rangle = x_{mn} + (\alpha_1 x_{1m}+\alpha_2 x_{2m})(\alpha_1 x_{1n}+\alpha_2 x2n) \qquad (39)$$

which is the simply the component form of the equation

$$\langle \mathbf{Y}\mathbf{Y}^T\rangle = [\mathbf{X}] + [\mathbf{X}]\boldsymbol{\alpha}\boldsymbol{\alpha}^T[\mathbf{X}] \qquad (40)$$

Using this in Eq. 29 gives

$$\begin{aligned}\langle \mathbf{a}\mathbf{a}^T\rangle &= [\mathbf{X}]^{-1}\left([\mathbf{X}] + [\mathbf{X}]\boldsymbol{\alpha}\boldsymbol{\alpha}^T[\mathbf{X}]\right)[\mathbf{X}]^{-1} && (41)\\ &= [\mathbf{X}]^{-1} + \boldsymbol{\alpha}\boldsymbol{\alpha}^T && (42)\end{aligned}$$

Finally, using this in Eq. 26 gives

$$[\boldsymbol{\sigma}_a^2] = [\mathbf{X}]^{-1} \qquad (43)$$

completing the proof.