Two supplements :

(1) Population version of ANOVA.

For completeness, we derive the ANOVA identity here for population version:

$$cov(\mathbf{x}) = cov(E(\mathbf{x}|Y)) + E(cov(\mathbf{x}|Y))$$

Here $\mathbf{x}$ is a p-dimensional random vector and Y is a random variable (or vector).

First write

$$\mathbf{x} = E(\mathbf{x}|Y) + (\mathbf{x} - E(\mathbf{x}|Y))$$

Because $E[\mathbf{x} - E(\mathbf{x}|Y)] = 0$, it follows that $cov(E(\mathbf{x}|Y), \mathbf{x} - E(\mathbf{x}|Y))$=$E[E(\mathbf{x}|Y)(\mathbf{x} - E(\mathbf{x}|Y))'] = E(E[E(\mathbf{x}|Y)(\mathbf{x}-E(\mathbf{x}|Y)')|Y]) = E(E(\mathbf{x}|Y)[E((\mathbf{x}-E(\mathbf{x}|Y))|Y)]) = E[(E(\mathbf{x}|Y)0] = 0$.

Thus we have $cov(\mathbf{x}) = cov(E(\mathbf{x}|Y)) + cov(\mathbf{x} - E(x|Y))$ Now

$$cov(\mathbf{x} - E(\mathbf{x}|Y)) = E(\mathbf{x} - E(\mathbf{x}|Y))(\mathbf{x} - E(\mathbf{x}|Y))')$$

$$= E(E[(\mathbf{x} - E(\mathbf{x}|Y))(\mathbf{x} - E(\mathbf{x}|Y))'|Y]) = E(cov(\mathbf{x}|Y))$$

We have derived the ANOVA identity.

(2) Finding $E(\mathbf{x}|\beta'\mathbf{x})$, under $\mathbf{x} = 0$ and the conditional linearity assumption that for any vector b, there exists a constant c such that

$$E(b'\mathbf{x}|\beta'\mathbf{x}) = c\beta'\mathbf{x}$$

(note c may depend on b).

First suppose $cov(\mathbf{x}) = I$. Then for any vector b, $cov(b'\mathbf{x}, \beta'\mathbf{x}) = b'cov(\mathbf{x})\beta = b'\beta$.

Now, for any two random variables with mean zero, say V, W, it is clear that $cov(W, V) = E(VW) = E(E(VW|V)) = E(VE(W|V))$. Taking $V = \beta'\mathbf{x}$ and $W = b'\mathbf{x}$, we see that $cov(b'\mathbf{x}, \beta'\mathbf{x}) = E(\beta'\mathbf{x}E(b'\mathbf{x}|\beta'\mathbf{x}))$, which , due to the linearity condition, equals to $E(\beta'\mathbf{x}c\beta'\mathbf{x}) = cvar(\beta'\mathbf{x}) = c\beta'\beta$

Therefore, we have shown that $c = b'\beta(\beta'\beta)^{-1}$ Since this is true for any b, it must hold that $E(\mathbf{x}|\beta'\mathbf{x}) = \beta(\beta'\beta)^{-1}\beta'\mathbf{x}$ We have seen that $E(\mathbf{x}|\beta'\mathbf{x})$ must be proportional to β.

For the general covariance matrix of $\mathbf{x}$, we can obtain $b'\Sigma_{\mathbf{x}}\beta = c\beta'\Sigma_{\mathbf{x}}\beta$. Thus $E(\mathbf{x}|Y) = \Sigma_{\mathbf{x}}\beta(\beta'\Sigma_{\mathbf{x}}\beta)^{-1}\beta'\mathbf{x}$, which is proportional to $\Sigma_{\mathbf{x}}\beta$.