6.10 Derive the relationship in Section 6.5.2 in two ways:

a. By completing the square

b. By expressing the Gaussian in its exponential family form

The product of two Gaussians $\mathcal{N}(\mathbf{x}|\mathbf{a},\mathbf{A})\ \mathcal{N}(\mathbf{x}|\mathbf{b},\mathbf{B})$ is an unnormalized Gaussian distribution $c\mathcal{N}(\mathbf{x}|\mathbf{c},\mathbf{C})$ with

$$\mathbf{C} = (\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}$$

$$\mathbf{c} = \mathbf{C}(\mathbf{A}^{-1}\mathbf{a} + \mathbf{B}^{-1}\mathbf{b})$$

$$c = (2\pi)^{-\frac{D}{2}}|\mathbf{A}+\mathbf{B}|^{-\frac{1}{2}} exp\left(-\frac{1}{2}(\mathbf{a}-\mathbf{b})^T(\mathbf{A}+\mathbf{B})^{-1}(\mathbf{a}-\mathbf{b})\right).$$

Note that the normalizing constant c itself can be considered a (normalized) Gaussian distribution either in $\mathbf{a}$ or in $\mathbf{b}$ with an "inflated" covariance matrix $\mathbf{A}+\mathbf{B}$, i.e., $c = \mathcal{N}(\mathbf{a}|\mathbf{b},\mathbf{A}+\mathbf{B}) = \mathcal{N}(\mathbf{b}|\mathbf{a},\mathbf{A}+\mathbf{B})$.

Solution.

a. By completing the square

$\mathcal{N}(\mathbf{x}|\mathbf{a},\mathbf{A}) = (2\pi)^{-\frac{D}{2}}|\mathbf{A}|^{-\frac{1}{2}} exp[-\frac{1}{2}(\mathbf{x}-\mathbf{a})^T\mathbf{A}^{-1}(\mathbf{x}-\mathbf{a})]$

$\mathcal{N}(\mathbf{x}|\mathbf{b},\mathbf{B}) = (2\pi)^{-\frac{D}{2}}|\mathbf{B}|^{-\frac{1}{2}} exp[-\frac{1}{2}(\mathbf{x}-\mathbf{b})^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{b})]$

$$\mathcal{N}(\mathbf{x}|\mathbf{a},\mathbf{A})\mathcal{N}(\mathbf{x}|\mathbf{b},\mathbf{B})$$

$$=$$

$$\left((2\pi)^{-\frac{D}{2}}|\mathbf{A}|^{-\frac{1}{2}} exp[-\frac{1}{2}(\mathbf{x}-\mathbf{a})^T\mathbf{A}^{-1}(\mathbf{x}-\mathbf{a})]\right)\left((2\pi)^{-\frac{D}{2}}|\mathbf{B}|^{-\frac{1}{2}} exp[-\frac{1}{2}(\mathbf{x}-\mathbf{b})^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{b})]\right)$$

The exponent term is

$-\frac{1}{2}(\mathbf{x}-\mathbf{a})^T\mathbf{A}^{-1}(\mathbf{x}-\mathbf{a}) - \frac{1}{2}(\mathbf{x}-\mathbf{b})^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{b})$

$= -\frac{1}{2}\left((\mathbf{x}-\mathbf{a})^T\mathbf{A}^{-1}(\mathbf{x}-\mathbf{a}) + (\mathbf{x}-\mathbf{b})^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{b})\right)$

$= -\frac{1}{2}(\mathbf{x}^T\mathbf{A}^{-1}\mathbf{x} - \mathbf{a}^T\mathbf{A}^{-1}\mathbf{x} - \mathbf{x}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{x}^T\mathbf{B}^{-1}\mathbf{x} - \mathbf{b}^T\mathbf{B}^{-1}\mathbf{x} - \mathbf{x}^T\mathbf{B}^{-1}\mathbf{b} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b})$

$= -\frac{1}{2}(\mathbf{x}^T(\mathbf{A}^{-1}+\mathbf{B}^{-1})\mathbf{x} - \mathbf{x}^T(\mathbf{A}^{-1}\mathbf{a}+\mathbf{B}^{-1}\mathbf{b}) - (\mathbf{a}^T\mathbf{A}^{-1}+\mathbf{b}^T\mathbf{B}^{-1})\mathbf{x} + \mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b})$

$(\mathbf{a}^T\mathbf{A}^{-1}+\mathbf{b}^T\mathbf{B}^{-1})\mathbf{x}$ is a scalar, so it is equal to its own transpose.

$\implies (\mathbf{a}^T\mathbf{A}^{-1}+\mathbf{b}^T\mathbf{B}^{-1})\mathbf{x} = ((\mathbf{a}^T\mathbf{A}^{-1}+\mathbf{b}^T\mathbf{B}^{-1})\mathbf{x})^T$

$= \mathbf{x}^T(\mathbf{a}^T\mathbf{A}^{-1}+\mathbf{b}^T\mathbf{B}^{-1})^T$

$= \mathbf{x}^T(\mathbf{A}^{-1}\mathbf{a}+\mathbf{B}^{-1}\mathbf{b})$ inverse of a symmertric matrix is also symmetric.

Substituting, we get

$= -\frac{1}{2}(\mathbf{x}^T(\mathbf{A}^{-1}+\mathbf{B}^{-1})\mathbf{x} - 2\mathbf{x}^T(\mathbf{A}^{-1}\mathbf{a}+\mathbf{B}^{-1}\mathbf{b}) + \mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b})$

$= -\frac{1}{2}(\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x} - 2\mathbf{x}^T\mathbf{C}^{-1}\mathbf{c} + \mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b})$

Adding and subtracting $\mathbf{c}^T\mathbf{C}^{-1}\mathbf{c}$,

$= -\frac{1}{2}(\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x} - 2\mathbf{x}^T\mathbf{C}^{-1}\mathbf{c} + \mathbf{c}^T\mathbf{C}^{-1}\mathbf{c} - \mathbf{c}^T\mathbf{C}^{-1}\mathbf{c} + \mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b})$

$= -\frac{1}{2}((\mathbf{x}-\mathbf{c})^T\mathbf{C}^{-1}(\mathbf{x}-\mathbf{c}) + \mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b} - \mathbf{c}^T\mathbf{C}^{-1}\mathbf{c})$

This got us the required term $(\mathbf{x}-\mathbf{c})^T\mathbf{C}^{-1}(\mathbf{x}-\mathbf{c})$ in the exponent.

Now we need to simplify $\mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b} - \mathbf{c}^T\mathbf{C}^{-1}\mathbf{c}$

$= \mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b} - (\mathbf{C}(\mathbf{A}^{-1}\mathbf{a}+\mathbf{B}^{-1}\mathbf{b}))^T\mathbf{C}^{-1}(\mathbf{C}(\mathbf{A}^{-1}\mathbf{a}+\mathbf{B}^{-1}\mathbf{b}))$

$= \mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b} - (\mathbf{C}(\mathbf{A}^{-1}\mathbf{a}+\mathbf{B}^{-1}\mathbf{b}))^T(\mathbf{A}^{-1}\mathbf{a}+\mathbf{B}^{-1}\mathbf{b})$

$= \mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b} - ((\mathbf{A}^{-1}\mathbf{a} + \mathbf{B}^{-1}\mathbf{b})^T\mathbf{C}^T)(\mathbf{A}^{-1}\mathbf{a} + \mathbf{B}^{-1}\mathbf{b})$

$= \mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b} - (\mathbf{a}^T\mathbf{A}^{-1} + \mathbf{b}^T\mathbf{B}^{-1})\mathbf{C}(\mathbf{A}^{-1}\mathbf{a} + \mathbf{B}^{-1}\mathbf{b})$

because covariance matrices are symmetric

$= \mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b} - (\mathbf{a}^T\mathbf{A}^{-1} + \mathbf{b}^T\mathbf{B}^{-1})(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}(\mathbf{A}^{-1}\mathbf{a} + \mathbf{B}^{-1}\mathbf{b})$

$= \mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b} - (\mathbf{a}^T\mathbf{A}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1} + \mathbf{b}^T\mathbf{B}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1})(\mathbf{A}^{-1}\mathbf{a} + \mathbf{B}^{-1}\mathbf{b})$

$= \mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b} - (\mathbf{a}^T\mathbf{A}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{A}^{-1}\mathbf{a} + \mathbf{a}^T\mathbf{A}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{B}^{-1}\mathbf{b} + \mathbf{b}^T\mathbf{B}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{B}^{-1}\mathbf{b})$

$= \mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b} - \mathbf{a}^T\mathbf{A}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{A}^{-1}\mathbf{a} - \mathbf{a}^T\mathbf{A}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{B}^{-1}\mathbf{b} - \mathbf{b}^T\mathbf{B}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{A}^{-1}\mathbf{a} - \mathbf{b}^T\mathbf{B}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{B}^{-1}\mathbf{b}$

$= \mathbf{a}^T(\mathbf{A}^{-1} - \mathbf{A}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{A}^{-1})\mathbf{a} + \mathbf{b}^T(\mathbf{B}^{-1} - \mathbf{B}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{B}^{-1})\mathbf{b} - \mathbf{a}^T\mathbf{A}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{B}^{-1}\mathbf{b} - \mathbf{b}^T\mathbf{B}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{A}^{-1}\mathbf{a}$

$= \mathbf{a}^T\mathbf{A}^{-1}(\mathbf{I} - (\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{A}^{-1})\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}(\mathbf{I} - (\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{B}^{-1})\mathbf{b} - \mathbf{a}^T\mathbf{A}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{B}^{-1}\mathbf{b} - \mathbf{b}^T\mathbf{B}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{A}^{-1}\mathbf{a}$

Using the property $(\mathbf{XY})^{-1} = \mathbf{Y}^{-1}\mathbf{X}^{-1}$ we obtain,

$\mathbf{A}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{B}^{-1} = (\mathbf{B}(\mathbf{A}^{-1} + \mathbf{B}^{-1})\mathbf{A})^{-1} = (\mathbf{B} + \mathbf{A})^{-1}$

$\Longrightarrow (\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{B}^{-1} = \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}$

Similarly, $\mathbf{B}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{A}^{-1} = (\mathbf{A}(\mathbf{A}^{-1} + \mathbf{B}^{-1})\mathbf{B})^{-1} = (\mathbf{A} + \mathbf{B})^{-1}$

$\Longrightarrow (\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{A}^{-1} = \mathbf{B}(\mathbf{A} + \mathbf{B})^{-1}$

Substituting above, we get

$= \mathbf{a}^T\mathbf{A}^{-1}(\mathbf{I} - \mathbf{B}(\mathbf{A} + \mathbf{B})^{-1})\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}(\mathbf{I} - \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1})\mathbf{b} - \mathbf{a}^T(\mathbf{B} + \mathbf{A})^{-1}\mathbf{b} - \mathbf{b}^T(\mathbf{B} + \mathbf{A})^{-1}\mathbf{a}$

$= \mathbf{a}^T\mathbf{A}^{-1}((\mathbf{A}+\mathbf{B})(\mathbf{A}+\mathbf{B})^{-1} - \mathbf{B}(\mathbf{A}+\mathbf{B})^{-1})\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}((\mathbf{A}+\mathbf{B})(\mathbf{A}+\mathbf{B})^{-1} - \mathbf{A}(\mathbf{A}+\mathbf{B})^{-1})\mathbf{b} - \mathbf{a}^T(\mathbf{B}+\mathbf{A})^{-1}\mathbf{b} - \mathbf{b}^T(\mathbf{B}+\mathbf{A})^{-1}\mathbf{a}$

$(\mathbf{A}+\mathbf{B})^{-1}$ exists because $\mathbf{A}$ and $\mathbf{B}$ are symmetric positive definite matrices, and therefore their sum is also a symmetric positive definite matrix, which makes $(\mathbf{A}+\mathbf{B})$ invertible.

$= \mathbf{a}^T\mathbf{A}^{-1}((\mathbf{A}+\mathbf{B}) - \mathbf{B})(\mathbf{A}+\mathbf{B})^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}((\mathbf{A}+\mathbf{B}) - \mathbf{A})(\mathbf{A}+\mathbf{B})^{-1}\mathbf{b} - \mathbf{a}^T(\mathbf{B}+\mathbf{A})^{-1}\mathbf{b} - \mathbf{b}^T(\mathbf{B}+\mathbf{A})^{-1}\mathbf{a}$

$= \mathbf{a}^T\mathbf{A}^{-1}(\mathbf{A})(\mathbf{A}+\mathbf{B})^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}(\mathbf{B})(\mathbf{A}+\mathbf{B})^{-1}\mathbf{b} - \mathbf{a}^T(\mathbf{B}+\mathbf{A})^{-1}\mathbf{b} - \mathbf{b}^T(\mathbf{B}+\mathbf{A})^{-1}\mathbf{a}$

$= \mathbf{a}^T(\mathbf{A}+\mathbf{B})^{-1}\mathbf{a} + \mathbf{b}^T(\mathbf{A}+\mathbf{B})^{-1}\mathbf{b} - \mathbf{a}^T(\mathbf{B}+\mathbf{A})^{-1}\mathbf{b} - \mathbf{b}^T(\mathbf{B}+\mathbf{A})^{-1}\mathbf{a}$

$= (\mathbf{a}-\mathbf{b})^T(\mathbf{A}+\mathbf{B})^{-1}(\mathbf{a}-\mathbf{b})$

Now, the constant term is

$(2\pi)^{-\frac{D}{2}}|\mathbf{A}|^{-\frac{1}{2}}(2\pi)^{-\frac{D}{2}}|\mathbf{B}|^{-\frac{1}{2}}$

We need $|\mathbf{C}|^{-\frac{1}{2}}$ here, so we multiply and divide by the same, and we get

$(2\pi)^{-\frac{D}{2}}|\mathbf{C}|^{-\frac{1}{2}}|\mathbf{C}|^{\frac{1}{2}}|\mathbf{A}|^{-\frac{1}{2}}(2\pi)^{-\frac{D}{2}}|\mathbf{B}|^{-\frac{1}{2}} = \left((2\pi)^{-\frac{D}{2}}|\mathbf{C}|^{-\frac{1}{2}}\right)\left((2\pi)^{-\frac{D}{2}}\frac{|\mathbf{A}|^{-\frac{1}{2}}|\mathbf{B}|^{-\frac{1}{2}}}{|\mathbf{C}|^{-\frac{1}{2}}}\right)$

We can use the properties that

a) the determinant of a matrix product is product of the determinants, and

b) determinant of a matrix inverse is the inverse of the determinant of this matrix, and write

$\frac{|\mathbf{A}|^{-\frac{1}{2}}|\mathbf{B}|^{-\frac{1}{2}}}{|\mathbf{C}|^{-\frac{1}{2}}} = (|\mathbf{A}||\mathbf{C}|^{-1}|\mathbf{B}|)^{-\frac{1}{2}} = |\mathbf{A}\mathbf{C}^{-1}\mathbf{B}|^{-\frac{1}{2}} = |\mathbf{A}(\mathbf{A}^{-1}+\mathbf{B}^{-1})\mathbf{B}|^{-\frac{1}{2}}$
$= |\mathbf{A}(\mathbf{A}^{-1}\mathbf{B}+\mathbf{I})|^{-\frac{1}{2}} = |(\mathbf{B}+\mathbf{A})|^{-\frac{1}{2}} = |(\mathbf{A}+\mathbf{B})|^{-\frac{1}{2}}$

Substituting, the constant term becomes

$$\left((2\pi)^{-\frac{D}{2}}|\mathbf{C}|^{-\frac{1}{2}}\right)\left((2\pi)^{-\frac{D}{2}}|(\mathbf{A}+\mathbf{B})|^{-\frac{1}{2}}\right)$$

Finally, putting it all together, the product of the two Gaussians becomes

$$\left((2\pi)^{-\frac{D}{2}}|\mathbf{C}|^{-\frac{1}{2}}\right)\left((2\pi)^{-\frac{D}{2}}|(\mathbf{A}+\mathbf{B})|^{-\frac{1}{2}}\right) exp\left[-\tfrac{1}{2}((\mathbf{x}-\mathbf{c})^T\mathbf{C}^{-1}(\mathbf{x}-\mathbf{c})+(\mathbf{a}-\mathbf{b})^T(\mathbf{A}+\mathbf{B})^{-1}(\mathbf{a}-\mathbf{b}))\right]$$

$$= \left((2\pi)^{-\frac{D}{2}}|(\mathbf{A}+\mathbf{B})|^{-\frac{1}{2}}\right) exp\left[-\tfrac{1}{2}((\mathbf{a}-\mathbf{b})^T(\mathbf{A}+\mathbf{B})^{-1}(\mathbf{a}-\mathbf{b}))\right]$$
$$\left((2\pi)^{-\frac{D}{2}}|\mathbf{C}|^{-\frac{1}{2}}\right) exp[-\tfrac{1}{2}((\mathbf{x}-\mathbf{c})^T\mathbf{C}^{-1}(\mathbf{x}-\mathbf{c}))]$$

$$= c\mathcal{N}(\mathbf{x}|\mathbf{c},\mathbf{C})$$

b. By expressing the Gaussian in its exponential family form

First, we have to find the exponential family form of the Gaussian.

$\mathcal{N}(\mathbf{x}|\mathbf{a},\mathbf{A}) = (2\pi)^{-\frac{D}{2}}|\mathbf{A}|^{-\frac{1}{2}}exp[-\frac{1}{2}(\mathbf{x}-\mathbf{a})^T\mathbf{A}^{-1}(\mathbf{x}-\mathbf{a})]$ needs to be expressed in the form $p(\mathbf{x}|\boldsymbol{\theta}) = h(\mathbf{x})exp(\langle\boldsymbol{\theta},\phi(\mathbf{x})\rangle - A(\boldsymbol{\theta}))$.

$$(2\pi)^{-\frac{D}{2}}|\mathbf{A}|^{-\frac{1}{2}}exp[-\tfrac{1}{2}(\mathbf{x}-\mathbf{a})^T\mathbf{A}^{-1}(\mathbf{x}-\mathbf{a})]$$
$$= (2\pi)^{-\frac{D}{2}}|\mathbf{A}|^{-\frac{1}{2}}exp[-\tfrac{1}{2}(\mathbf{x}^T\mathbf{A}^{-1}\mathbf{x}-2\mathbf{a}^T\mathbf{A}^{-1}\mathbf{x}+\mathbf{a}^T\mathbf{A}^{-1}\mathbf{a})]$$
$$= exp[ln((2\pi)^{-\frac{D}{2}}|\mathbf{A}|^{-\frac{1}{2}}) - \tfrac{1}{2}(\mathbf{x}^T\mathbf{A}^{-1}\mathbf{x}-2\mathbf{a}^T\mathbf{A}^{-1}\mathbf{x}+\mathbf{a}^T\mathbf{A}^{-1}\mathbf{a})]$$

$\mathbf{x}^T\mathbf{A}^{-1}\mathbf{x}$ can be written as this dot product : $\langle vec(\mathbf{x}\mathbf{x}^T), vec(\mathbf{A}^{-1})\rangle$

Let $\phi(\mathbf{x}) = \begin{bmatrix} vec(\mathbf{x}\mathbf{x}^T) \\ \mathbf{x} \end{bmatrix}$,

$\boldsymbol{\theta}_1 = -\frac{1}{2}\begin{bmatrix} vec(\mathbf{A}^{-1}) \\ -2\mathbf{A}^{-1}\mathbf{a} \end{bmatrix}$

$h(\mathbf{x}) = 1,$

$A(\boldsymbol{\theta}_1) = -(ln((2\pi)^{-\frac{D}{2}}|\mathbf{A}|^{-\frac{1}{2}}) - \frac{1}{2}\mathbf{a}^T\mathbf{A}^{-1}\mathbf{a})$

That gives us the exponential family representation of the Gaussian distribution, such that

$\mathcal{N}(\mathbf{x}|\mathbf{a}, \mathbf{A}) = h(\mathbf{x})exp(\langle\boldsymbol{\theta}_1, \phi(\mathbf{x})\rangle - A(\boldsymbol{\theta}_1)),$

$\mathcal{N}(\mathbf{x}|\mathbf{b}, \mathbf{B}) = h(\mathbf{x})exp(\langle\boldsymbol{\theta}_2, \phi(\mathbf{x})\rangle - A(\boldsymbol{\theta}_2)),$

where $\boldsymbol{\theta}_2 = -\frac{1}{2}\begin{bmatrix} vec(\mathbf{B}^{-1}) \\ -2\mathbf{B}^{-1}\mathbf{b} \end{bmatrix}$

and $A(\boldsymbol{\theta}_2) = -(ln((2\pi)^{-\frac{D}{2}}|\mathbf{B}|^{-\frac{1}{2}}) - \frac{1}{2}\mathbf{b}^T\mathbf{B}^{-1}\mathbf{b})$

The product of the two Gaussian distributions is given by:

$\mathcal{N}(\mathbf{x}|\mathbf{a}, \mathbf{A})\mathcal{N}(\mathbf{x}|\mathbf{b}, \mathbf{B}) = h(\mathbf{x})exp(\langle\boldsymbol{\theta}_1, \phi(\mathbf{x})\rangle - A(\boldsymbol{\theta}_1))h(\mathbf{x})exp(\langle\boldsymbol{\theta}_2, \phi(\mathbf{x})\rangle - A(\boldsymbol{\theta}_2))$

$= h(\mathbf{x})^2 exp(\langle\boldsymbol{\theta}_1, \phi(\mathbf{x})\rangle - A(\boldsymbol{\theta}_1) + \langle\boldsymbol{\theta}_2, \phi(\mathbf{x})\rangle - A(\boldsymbol{\theta}_2))$

$= exp(\langle(\boldsymbol{\theta}_1 + \boldsymbol{\theta}_2), \phi(\mathbf{x})\rangle - (A(\boldsymbol{\theta}_1) + A(\boldsymbol{\theta}_2)))$

Now we calculate the exponent term of $\mathcal{N}(\mathbf{x}|\mathbf{a}, \mathbf{A})\mathcal{N}(\mathbf{x}|\mathbf{b}, \mathbf{B})$

$\boldsymbol{\theta}_1 + \boldsymbol{\theta}_2 = -\frac{1}{2}\begin{bmatrix} vec(\mathbf{A}^{-1}) + vec(\mathbf{B}^{-1}) \\ -2\mathbf{A}^{-1}\mathbf{a} - 2\mathbf{B}^{-1}\mathbf{b} \end{bmatrix}$

$= -\frac{1}{2}\begin{bmatrix} vec(\mathbf{A}^{-1} + \mathbf{B}^{-1}) \\ -2(\mathbf{A}^{-1}\mathbf{a} + \mathbf{B}^{-1}\mathbf{b}) \end{bmatrix}$

$$= -\frac{1}{2}\begin{bmatrix} vec(\mathbf{C}^{-1}) \\ -2(\mathbf{C}^{-1}\mathbf{c}) \end{bmatrix}$$

$$\Longrightarrow \langle(\boldsymbol{\theta}_1 + \boldsymbol{\theta}_2), \phi(\mathbf{x})\rangle = -\frac{1}{2}\begin{bmatrix} vec(\mathbf{C}^{-1}) & -2(\mathbf{C}^{-1}\mathbf{c}) \end{bmatrix}\begin{bmatrix} vec(\mathbf{x}\mathbf{x}^T) \\ \mathbf{x} \end{bmatrix}$$

$$= -\frac{1}{2}\left(\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x} - 2\mathbf{x}^T\mathbf{C}^{-1}\mathbf{c}\right)$$

$$A(\boldsymbol{\theta}_1) + A(\boldsymbol{\theta}_2)$$

$$= ln((2\pi)^{-\frac{D}{2}}|\mathbf{A}|^{-\frac{1}{2}}) - \frac{1}{2}\mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + ln((2\pi)^{-\frac{D}{2}}|\mathbf{B}|^{-\frac{1}{2}}) - \frac{1}{2}\mathbf{b}^T\mathbf{B}^{-1}\mathbf{b}$$

$$= ln((2\pi)^{-D}|\mathbf{A}|^{-\frac{1}{2}}|\mathbf{B}|^{-\frac{1}{2}}) - \frac{1}{2}\left(\mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b}\right)$$

From the results we have already obtained in part a., we know that

$$\mathbf{a}^T\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b} = (\mathbf{a} - \mathbf{b})^T(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{a} - \mathbf{b}) + \mathbf{c}^T\mathbf{C}^{-1}\mathbf{c}$$

and $\frac{|\mathbf{A}|^{-\frac{1}{2}}|\mathbf{B}|^{-\frac{1}{2}}}{|\mathbf{C}|^{-\frac{1}{2}}} = |(\mathbf{A} + \mathbf{B})|^{-\frac{1}{2}} \Longrightarrow |\mathbf{A}|^{-\frac{1}{2}}|\mathbf{B}|^{-\frac{1}{2}} = |(\mathbf{A} + \mathbf{B})|^{-\frac{1}{2}}|\mathbf{C}|^{-\frac{1}{2}}$

Substituting, we get

$$A(\boldsymbol{\theta}_1) + A(\boldsymbol{\theta}_2) = ln((2\pi)^{-D}(|(\mathbf{A}+\mathbf{B})|^{-\frac{1}{2}}|\mathbf{C}|^{-\frac{1}{2}})) - \frac{1}{2}\left((\mathbf{a} - \mathbf{b})^T(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{a} - \mathbf{b}) + \mathbf{c}^T\mathbf{C}^{-1}\mathbf{c}\right)$$

The full exponent term becomes :

$$ln((2\pi)^{-D}(|(\mathbf{A}+\mathbf{B})|^{-\frac{1}{2}}|\mathbf{C}|^{-\frac{1}{2}})) - \frac{1}{2}\left((\mathbf{a} - \mathbf{b})^T(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{a} - \mathbf{b}) + \mathbf{c}^T\mathbf{C}^{-1}\mathbf{c} + \mathbf{x}^T\mathbf{C}^{-1}\mathbf{x} - 2\mathbf{x}^T\mathbf{C}^{-1}\mathbf{c}\right)$$

$$= ln((2\pi)^{-D}(|(\mathbf{A}+\mathbf{B})|^{-\frac{1}{2}}|\mathbf{C}|^{-\frac{1}{2}})) - \frac{1}{2}\left((\mathbf{a} - \mathbf{b})^T(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{a} - \mathbf{b}) + (\mathbf{x}\text{-}\mathbf{c})^T\mathbf{C}^{-1}(\mathbf{x}\text{-}\mathbf{c})\right)$$

Taking the exponent, it becomes

$$exp\left(ln((2\pi)^{-D}(|(\mathbf{A} + \mathbf{B})|^{-\frac{1}{2}}|\mathbf{C}|^{-\frac{1}{2}})) - \frac{1}{2}\left((\mathbf{a} - \mathbf{b})^T(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{a} - \mathbf{b}) + (\mathbf{x}\text{-}\mathbf{c})^T\mathbf{C}^{-1}(\mathbf{x}\text{-}\mathbf{c})\right)\right)$$

$$= (2\pi)^{-D}(|(\mathbf{A}+\mathbf{B})|^{-\frac{1}{2}}|\mathbf{C}|^{-\frac{1}{2}})exp\left(-\frac{1}{2}\left((\mathbf{a} - \mathbf{b})^T(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{a} - \mathbf{b}) + (\mathbf{x}\text{-}\mathbf{c})^T\mathbf{C}^{-1}(\mathbf{x}\text{-}\mathbf{c})\right)\right)$$

$$= (2\pi)^{-\frac{D}{2}}(2\pi)^{-\frac{D}{2}}|(\mathbf{A}+\mathbf{B})|^{-\frac{1}{2}}|\mathbf{C}|^{-\frac{1}{2}}exp\left(-\frac{1}{2}(\mathbf{a} - \mathbf{b})^T(\mathbf{A} + \mathbf{B})^{-1}(\mathbf{a} - \mathbf{b})\right)exp\left(-\frac{1}{2}(\mathbf{x}\text{-}\mathbf{c})^T\mathbf{C}^{-1}(\mathbf{x}\text{-}\mathbf{c})\right)$$

$= (2\pi)^{-\frac{D}{2}}|(\mathbf{A}+\mathbf{B})|^{-\frac{1}{2}}exp\left(-\frac{1}{2}(\mathbf{a}-\mathbf{b})^T(\mathbf{A}+\mathbf{B})^{-1}(\mathbf{a}-\mathbf{b})\right)(2\pi)^{-\frac{D}{2}}|\mathbf{C}|^{-\frac{1}{2}}exp\left(-\frac{1}{2}(\mathbf{x}\text{-}\mathbf{c})^T\mathbf{C}^{-1}(\mathbf{x}\text{-}\mathbf{c})\right)$

$= c\mathcal{N}(\mathbf{x}|\mathbf{c},\mathbf{C})$

Hence proved.