A2. E1.

$\because \langle \cdot , \cdot \rangle$ is an inner product.

$\therefore \langle \cdot , \cdot \rangle$ is positive defined, linear and bilinear.

$\therefore \forall x \in V \setminus \{0\}$, $\langle x, x \rangle > 0$. $\langle 0, 0 \rangle = 0$.

$\therefore \forall x \in V \setminus \{0\}$, $\|x\| = \sqrt{\langle x, x \rangle} > 0$. $\|0\| = 0$.

$\therefore \|\cdot\|$ is positive defined.

For $\lambda \in \mathbb{R}$, $\because \langle \cdot , \cdot \rangle$ is linear,

$\therefore \|\lambda x\| = \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{\lambda^2 \langle x, x \rangle} = |\lambda| \sqrt{\langle x, x \rangle} = |\lambda| \, \|x\|$.

$\therefore \|\cdot\|$ is absolutely homogeneous.

$\|x+y\| = \sqrt{\langle x+y, x+y \rangle} = \sqrt{\langle x, x+y \rangle + \langle y, x+y \rangle} = \sqrt{\langle x, x \rangle + \langle y, y \rangle + 2 \langle x, y \rangle}$

$\|x\| + \|y\| = \sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle}$

$\|x+y\|^2 = \langle x, x \rangle + \langle y, y \rangle + 2 \langle x, y \rangle$

$(\|x\| + \|y\|)^2 = \langle x, x \rangle + \langle y, y \rangle + 2 \|x\| \, \|y\|$

According to Cauchy-Schwartz inequality, $\langle x, y \rangle \leq \|x\| \, \|y\|$

$\therefore \langle x, x \rangle + \langle y, y \rangle + 2 \langle x, y \rangle \leq \langle x, x \rangle + \langle y, y \rangle + 2 \|x\| \, \|y\|$.

$\therefore \|x+y\|^2 \leq (\|x\| + \|y\|)^2$

$\therefore \|x+y\| \leq \|x\| + \|y\|$.

$\therefore \|\cdot\|$ is trangle inequality.

$\therefore \|\cdot\|$ is a norm.

A2. E2. 1.

Let $f(x) = x^T$, $g(x) = ab^T x$. $\therefore x^T a b^T x = f(x) g(x)$

$\therefore \nabla_x (x^T a b^T x) = \nabla_x (f(x) g(x))$

$$= \frac{\partial f}{\partial x} g(x) + f(x) \frac{\partial g}{\partial x}$$

$$= (a b^T x)^T + x^T a b^T$$

$$= x^T b a^T + x^T a b^T$$

$$= \langle x, b \rangle a^T + \langle x, a \rangle b^T$$

$$= \langle b, x \rangle a^T + \langle a, x \rangle b^T$$

$$= b^T x a^T + a^T x b^T$$

$$= a^T x b^T + b^T x a^T.$$

A2. E2. 2.

Let $f(x) = x^T$, $g(x) = Bx$. $\therefore x^T B x = f(x) g(x)$

$\therefore \nabla_x x^T B x = \nabla_x (f(x) g(x))$

$$= \frac{\partial f}{\partial x} g(x) + f(x) \frac{\partial g}{\partial x}$$

$$= (Bx)^T + x^T B$$

$$= x^T B^T + x^T B$$

$$= x^T (B^T + B)$$

$$= x^T (B + B^T)$$

A2. E3

Suppose $A, B \in V$.

$\because$ A, B are symmetric positive definite,

$\therefore$ $\forall x \in V \setminus \{0\}$, $x^T A x > 0$, $x^T B x > 0$.

$\therefore$ $\forall x \in V \setminus \{0\}$, $x^T (pA + qB) x$

$$= x^T pAx + x^T qBx$$

$$= p x^T A x + q x^T B x > 0.$$

$\therefore$ $pA + qB$ is positive definite.

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix} = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{1n} & \cdots & A_{nn} \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & \cdots & B_{1n} \\ \vdots & \ddots & \vdots \\ B_{n1} & \cdots & B_{nn} \end{bmatrix} = \begin{bmatrix} B_{11} & \cdots & B_{1n} \\ \vdots & \ddots & \vdots \\ B_{1n} & \cdots & B_{nn} \end{bmatrix}$$

$$pA + qB = p \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix} + q \begin{bmatrix} B_{11} & \cdots & B_{1n} \\ \vdots & \ddots & \vdots \\ B_{n1} & \cdots & B_{nn} \end{bmatrix}$$

$$= p \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{1n} & \cdots & A_{nn} \end{bmatrix} + q \begin{bmatrix} B_{11} & \cdots & B_{1n} \\ \vdots & \ddots & \vdots \\ B_{1n} & \cdots & B_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} pA_{11} + qB_{11} & \cdots & pA_{1n} + qB_{1n} \\ \vdots & \ddots & \vdots \\ pA_{1n} + qB_{1n} & \cdots & pA_{nn} + qB_{nn} \end{bmatrix}$$

$\therefore$ $pA + qB$ is symmetric.

$\therefore$ $pA + qB$ is symmetric and positive definite.

A2. E4. 1.

Let $f(\theta) = y - X\theta - c$. $g(f) = f^T A f$.

$\therefore \mathcal{L} = g(f(\theta)) + \theta^T B \theta + \|c\|_A^2$

$$\nabla_\theta \mathcal{L}(\theta, c) = \frac{\partial g}{\partial f} \frac{\partial f}{\partial \theta} + \frac{\partial}{\partial \theta} \theta^T B \theta.$$

$$= 2 f(\theta)^T A \frac{\partial f}{\partial \theta} + 2\theta^T B$$

$$= 2(y - X\theta - c)^T A \frac{\partial}{\partial \theta}(y - X\theta - c) + 2\theta^T B$$

$$= -2(y - X\theta - c)^T A X + 2\theta^T B$$

$$= 2\theta^T B - 2(y^T - \theta^T X^T - c^T) A X$$

$$= 2\theta^T B - 2y^T A X + 2\theta^T X^T A X + 2c^T A X$$

$$= 2\theta^T (B + X^T A X) - 2(y^T - c^T) A X$$

A2. E4.2.

$$\nabla_\theta \mathcal{L}(\theta, c) = 2\theta^T (B + X^T A X) - 2(y^T - c^T) A X = 0.$$

$$\theta^T (B + X^T A X) = (y^T - c^T) A X$$

$$(B + X^T A X)^T \theta = [(y^T - c^T) A X]^T$$

$$(B^T + X^T A X)\theta = [(y^T - c^T) A X]^T$$

$$(B + X^T A X)\theta = [(y^T - c^T) A X]^T$$

Suppose $B + X^T A X$ is not invertable. Let $v \in \mathbb{R}^D$.

$(B + X^T A X) v = 0$, $v \neq 0$.

$X^T A X = -B$

$\because$ A B is positive definite, X is full rank.

When $D = 1$, $B > 0$, $-B < 0$. But $X^T A X > 0$.

$\therefore X^T A X \neq -B$

$\therefore B + X^T A X$ is invertable. $\therefore \theta = (B + X^T A X)^{-1} [(y^T - c^T) A X]^T$

A2. E4. 3.

Let $f(c) = y - X\theta - c$. $g(f) = f^T A f$.

$$\therefore \mathcal{L}(\theta, c) = (y - X\theta - c)^T A (y - X\theta - c) + \|\theta\|_B^2 + c^T A c$$
$$= g(f(c)) + \|\theta\|^2_B + c^T A c.$$

$$\nabla_c \mathcal{L}(\theta, c) = \frac{\partial g}{\partial f} \frac{\partial f}{\partial c} + \frac{\partial}{\partial c} c^T A c$$
$$= 2 f^T(c) A \frac{\partial}{\partial c} (y - X\theta - c) + 2 c^T A$$
$$= -2 (y - X\theta - c)^T A + 2 c^T A$$
$$= -2 (y - X\theta)^T A + 4 c^T A$$

A2. E4. 4.

$\nabla_c \mathcal{L}(\theta, c) = 0.$

$-2 (y - X\theta)^T A + 4 c^T A = 0.$

$2 c^T A = (y - X\theta)^T A$

$2 A^T c = A^T (y - X\theta)$

A is symmetric positive definite,

$\therefore 2 A c = A (y - X\theta)$

$2 A^{-1} A c = A^{-1} A (y - X\theta)$

$c = \frac{1}{2} (y - X\theta).$

A2. E4. 5.

$A = I, c = 0, B = \lambda I,$

$$\theta = (\lambda I + X^T I X)^{-1} [(y^T - 0^T) I X]^T$$
$$= (\lambda I + X^T X)^{-1} (y^T X)^T$$
$$= (X^T X + \lambda I)^{-1} X^T y$$