1. To show K is a field:
We can say that $a+bi$ and $c+di$ are elements of K, as a and b are rational numbers and so $a+c$ and $b+d$ are rational and so $(a+c)+(b+d)i$ is an element of K.
Furthermore, $(ac-bd)$ and $(ad+bc)$ are rational numbers, so $(ac-bd)+(ad+bc)i$ is an element of K.
Finally, there exists additive and multiplicative inverses, as $-a-bi$ exists. Additionally, we know that the aforementioned identities exist as well.

2. To show that $U \cap W$ and $U+W$ are subspaces of V:
we know based on the description of each that they are non-empty subspaces of V. Additionally, they should contain the zero vector since U and W are subspaces of V

Considering x_1, x_2 to be elements of $U \cap W$, we can say that x_1 and x_2 belong to U and W. $x_1 + x_2$ belongs to both U and W, and thus, $U \cap W$.
Similarly, u_1 and u_2 being elements of U and w_1 and w_2 being elements of W would be written as $(u_1+w_1)+(u_2+w_2)$, which belongs to $U+W$.
Alternatively, we can let k be a scalar and x be an element of U and thus, kx belongs to $U \cap W$. Also, $ku + kw = k(u+w)$ belongs to $U+W$ as U and W are subspaces of V.
Thus, we have proved that $U \cap W$ and $U+W$ are subspaces of V.

3. • $\|A\| > 0$ if $A \neq 0$:
- Proof: If $A \neq 0$, then there needs to be at least one element a_{ij} in A that is nonzero. Thus, $\max_j \sum_{i=1}^{m} |a_{ij}| > 0$ and $\|A\| > 0$.

• $\|\gamma A\| = |\gamma| \cdot \|A\|$ for any scalar γ:
- Proof: Considering $X = \gamma A$, $\|X\| = \max_j \sum_{i=1}^{m} |\gamma a_{ij}| = |\gamma| \max_j \sum_{i=1}^{m} |a_{ij}| = |\gamma| \|A\|$
Thus, $\|\gamma A\| = \|X\| = |\gamma| \|A\|$

• $\|A+B\| \leq \|A\| + \|B\|$:
$$\max_j \sum_{i=1}^{m} |a_{ij} + b_{ij}| \leq \max_j \left(\sum_{i=1}^{m} |a_{ij}| + \sum_{i=1}^{m} |b_{ij}|\right) \leq \max_j \sum_{i=1}^{m} |a_{ij}| + \max_j \sum_{i=1}^{m} |b_{ij}| = \|A\| + \|B\|$$
Thus, $\|A+B\| \leq \|A\| + \|B\|$. Similarly, we can show that $\|A+B\|_\infty \leq \|A\|_\infty + \|B\|_\infty$.

• $\|AB\| \leq \|A\| \cdot \|B\|$:
- Proof: Considering $C = AB$, $|c_{ij}| = \left|\sum_{k=1}^{nm} a_{ik} b_{kj}\right| \leq \sum_{k=1}^{nm} |a_{ik}| \cdot |b_{jk}| = \max_k \sum_{k=1}^{nm} |a_{ik}| \cdot \max_k \sum_{k=1}^{nm} |b_{jk}|$
$= \|A\| \cdot \|B\|_\infty$

• $\|Ax\| \leq \|A\| \cdot \|x\|$ for any vector x:
- Proof: $\|Ax\| \leq \|A\| \cdot \|x\|$ iff $\|Ax\| / \|x\| \leq \|A\|$.
Considering $y = Ax$, then $\|y\| / \|x\| = \|Ax\| / \|x\| \leq \max_j \left(\sum_{i=1}^{m} |a_{ij}|\right) = \|A\|$
Thus, $\|Ax\| \leq \|A\| \cdot \|x\|$. (For $\|.\|$)

Thus, $||Ax|| \leq ||A|| \cdot ||x||$. (For $||.||$)

4. a) With x being a column vector:

$$x^T(Bx) = x^T\left(\frac{(B+B^T)}{2} + \frac{(B-B^T)}{2}\right)x$$
$$= \frac{x^T(B+B^T)}{2x} + \frac{x^T(B-B^T)}{2x}$$

- The first term is always going to be non-negative for any vector x, so the second term will be 0.

- B_s being the symmetric part of B, $B_s = \frac{(B+B^T)}{2}$. $B-B^T = 2(B-B_s)$, so we show that the quadratic form $x^T(B-B_s)x$ is 0, because $(B-B_s)=0$ and annihilates due to skew-symmetricism.

b) Considering A to be a square matrix, $B = A^TA$ is symmetric. For any vector x, we have:

$$x^T(A^TA)x = (Ax)^T(Ax) = ||Ax||^2$$

- $B = A^TA$ is non-negative definite because of iff $||Ax||^2 \geq 0$ for all $x \neq 0$, which is true because of the norm of any vector being non-negative.

5. The least-squares orthogonality principle states that the residual vector $b - A\hat{x}$ is orthogonal to the column space of A. Rewriting in terms of the SVD of A, we get $V(SD^Tb - DC\hat{x}) = 0$.
We can use the SVD to isolate $\hat{x}$:

$\hat{x} = VD^{-1}C^Tb$

Use this value in the least-squares orthogonality principle:

$V(SD^Tb - DCVD^{-1}C^Tb) = 0$

$\cdot V^T \qquad\qquad \cdot V^T$

$SD^Tb - DCVD^{-1}C^Tb = 0$

$\boxed{Db - C^{(1+T)}Db = 0}$

6. $f(x) = -\sum_{i=1}^{N} T_i x_i + \sum_{i=1}^{N} x_i \log(x_i)$

subject to constraint $\sum_{i=1}^{N} x_i = 1$

We define the Lagrangian as: $L(x,\lambda) = f(x) + \lambda\left(\sum_{i=1}^{N} x_i - 1\right)$

$\frac{\partial L}{\partial x_i} = -T_i + 1 + \log(x_i) + \lambda = 0$

$\frac{\partial L}{\partial \lambda} = \sum_{i=1}^{N} x_i - 1 = 0$

$$\frac{\partial L}{\partial x_i} = -T_i + 1 + \log(x_i) + \lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = \sum_{i=1}^{N} x_i - 1 = 0$$

We solve for x_i:

$$x_i = \exp(-T_i - 1 + \lambda)$$

Noting that the sum of x_i is 1:

$$\frac{\sum_{i=1}^{N} \exp(-T_i - 1 + \lambda)}{\sum_{j=1}^{N} \exp(-T_j - 1 + \lambda)} = 1$$

$$\ln\left[\sum_{i=1}^{N} \exp(-T_i - 1 + \lambda)\right] = \ln\left[\sum_{j=1}^{N} \exp(-T_j - 1 + \lambda)\right]$$

The methods of bisection and Newton's can be used to solve for λ.