1. (a) $f(\vec{x}) = 2x_1^2 + x_1 x_2 + x_2^2 - 3x_1 - 5x_2$

$$
\begin{aligned}
4f(\vec{x}) &= 8x_1^2 + 4x_1x_2 + 4x_2^2 - 12x_1 - 20x_2 \\
&\geq 6x_1^2 + 2x_2^2 - 12x_1 - 20x_2 \\
&= 6(x_1-1)^2 + 2(x_2-5)^2 - 56 \\
&\geq 2(x_1-1)^2 + 2(x_2-5)^2 - 56 \\
&\geq x_1^2 + x_2^2 - 108
\end{aligned}
$$

Obviously, when $\|\vec{x}\| \to \infty$, $f(\vec{x}) \to +\infty$, so $f(\vec{x})$ is coercive.

(b) $f(\vec{x})$ is continuous on $\mathbb{R}^2$. Given the conclusion in (a), we know that $f(\vec{x})$ has global min but does not have global max.

The global mins satisfy

$\nabla f(\vec{x}) = \vec{0}$, i.e. $\begin{cases} 4x_1 + x_2 - 3 = 0 \\ x_1 + 2x_2 - 5 = 0 \end{cases}$

$\begin{cases} x_1 = \frac{1}{7} \\ x_2 = \frac{17}{7} \end{cases}$ is the only solution, so it must be the global min

Thus, the global min of $f(\vec{x})$ is $\left(\frac{1}{7}, \frac{17}{7}\right)$

2. (a) When $\alpha \in \mathbb{R} \to +\infty$, $f(\alpha \vec{w_0}) \to 0$, so $\forall \varepsilon > 0$, $\exists A > 0$, $\forall \alpha > A$, $|f(\alpha \vec{w_0})| < \varepsilon$.

Since $f(\vec{w}) > 0$, $f(\vec{w})$ cannot have a global min.

(b) (i) For a given i,
$$f(\vec{w}) \geq \log(1 + e^{-y_i \vec{x_i}^T \vec{w}}) \geq \log e^{-y_i \vec{x_i}^T \vec{w}} = -y_i \vec{x_i}^T \vec{w}$$

Thus, $f(\vec{w}) \geq h(\vec{w})$

(ii) $h(\vec{w})$ is continuous on S, which is a compact set, so $h(\vec{w})$ has a global min $\vec{w_0}$ on S.

If $h(\vec{w_0}) \leq 0$, then $\forall i$, $-y_i \vec{x_i}^T \vec{w} \leq 0$, which contradicts the fact that the dataset is not linearly separable. Thus, $C \triangleq h(\vec{w_0}) > 0$

(iii) Let $\vec{w_1} = \frac{\vec{w}}{\|\vec{w}\|}$. Then $\vec{w_1} \in S$.

$$h(\vec{w}) = h(\|\vec{w}\| \vec{w_1}) = \|\vec{w}\| h(\vec{w_1}) \geq C\|\vec{w}\|.$$

(iv) $f(\vec{w}) \geq C\|\vec{w}\|$, so $f(\vec{w})$ is coercive. We also know that $f(\vec{w})$ is continuous, so $f(\vec{w})$ has a global min.

(c) Let $g(x) = \log(1 + e^x)$, $h_i(\vec{w}) = -y_i \vec{x_i}^T \vec{w}$.

Then $\left[g(h_i(\vec{w}))\right]' = g'(h_i(\vec{w})) \cdot h_i'(\vec{w})$

$$= -\frac{y_i e^{-y_i \vec{x_i}^T \vec{w}}}{e^{-y_i \vec{x_i}^T \vec{w}} + 1} \vec{x_i}^T$$

Thus, $\nabla f(\vec{w}) = f'(\vec{w})^T = -\sum_{i=1}^{m} \frac{y_i e^{-y_i \vec{x_i}^T \vec{w}}}{e^{-y_i \vec{x_i}^T \vec{w}} + 1} \vec{x_i}$

3. (a) We already know that

$$g(a+t) = g(a) + g'(a)t + \frac{1}{2}g''(a+\theta t)t^2, \ \theta \in (0,1)$$

Let $\hat{d} = \frac{\vec{d}}{\|\vec{d}\|}$.

Applying the formula to $g(0+t) = f(\vec{x} + t\hat{d})$ will obtain

$$f(\vec{x} + t\hat{d}) = g(0) + g'(0)t + \frac{1}{2}g''(\theta t)t^2$$

Because $g'(t) = f'(\vec{x} + t\hat{d})\hat{d}$, $g''(t) = \hat{d}^T \nabla^2 f(\vec{x} + t\hat{d})\hat{d}$,

when $t = \|\vec{d}\|$ we have:

$$f(\vec{x} + \vec{d}) = f(\vec{x}) + \nabla f(\vec{x} + \vec{d})\vec{d} + \frac{1}{2}\vec{d}^T \nabla^2 f(\vec{x} + \theta\vec{d})\vec{d},$$

for some $\theta \in (0,1)$

(b) Let $\vec{g}(t) = \nabla f(\vec{x} + t\vec{d})$. Then $\vec{g}'(t) = \nabla^2 f(\vec{x} + t\vec{d}) \cdot \vec{d}$

$$\int_0^1 \vec{g}'(t)\,dt = \vec{g}(1) - \vec{g}(0) = \nabla f(\vec{x} + \vec{d}) - \nabla f(\vec{x})$$

Thus, $\nabla f(\vec{x} + \vec{d}) = \nabla f(\vec{x}) + \int_0^1 \nabla^2 f(\vec{x} + t\vec{d})\vec{d}\,dt$

4. $A = \begin{pmatrix} 6 & 2 & 0 \\ 2 & 5 & -2 \\ 0 & -2 & 4 \end{pmatrix}$

$6 > 0$, $\begin{vmatrix} 6 & 2 \\ 2 & 5 \end{vmatrix} = 26 > 0$, $|A| = 80 > 0$,

so A is positive definite

$B = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 2 & 1 \\ 0 & 1 & -3 \end{pmatrix}$

$1 > 0$, $\begin{vmatrix} 1 & 2 \\ 2 & 2 \end{vmatrix} = -2 < 0$, so B is indefinite.

$C = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$

$2 > 0$, $\begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 2 > 0$, $|C| = 0$

All principal minors are non-negative,
so C is positive semidefinite.