9.21

1.(a) $f(\vec{x}) = (x_1-x_2)^2 + (\frac{1}{2}x_1 - 1)^2 + (x_2-1)^2 + \frac{3}{4}||x||^2 + \frac{1}{4}x_2^2 - 2$

it's obvious $f(\vec{x}) \to +\infty$ as $||\vec{x}|| \to +\infty$, $f(\vec{x})$ is coercive.

$f(\vec{x})$ is surely continuous.

So $f(\vec{x})$ has a global minimum

(b) $f(\vec{x}) = \frac{1}{2}||\vec{x}||^2 - 1 + (\frac{\sqrt{2}}{2}x_1 + \sqrt{2}x_2 - \frac{\sqrt{2}}{2})^2 + \frac{x_2^2}{2} + \frac{1}{2}$

continuous and $f(\vec{x}) \to +\infty$ as $||\vec{x}|| \to +\infty$

$f(x)$ has a global minimum. but I haven't figured out about hint. ($||\vec{x}||^2 - 1$?)

(c) on the line $x_1 + x_2 = 0$, $f(\vec{x}) = -x_1 - 2x_2 = x_1$

$f(\vec{x}) \to -\infty$ as $x_1 \to -\infty$

$f(x)$ has no global minimum

2.(a) $\nabla f(\vec{x}) = (\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \cdots \frac{\partial f}{\partial x_n})^T$

$= (x_1, x_2, x_3 \cdots)^T$

(b) $\nabla f(\vec{x}) = \nabla \frac{1}{2}||X\vec{w} - \vec{y}||^2 + \nabla \frac{1}{2}\lambda ||w||^2$

$= X^T \nabla_{X\vec{w}-\vec{y}} \frac{1}{2} ||X\vec{w} - \vec{y}|| + \nabla \frac{1}{2} \lambda w^T w$

$= X^T \cdot (X\vec{w} - \vec{y}) + \lambda$

not familiar with such derivation (vet attached)

3. (1) for the specific w_0, $y_i x_i^T w_0 > 0$ for $\forall i = 1, 2, \cdots, m$

we have nw_0 s.t. $y_i x_i^T (nw_0) > 0$

$f(w) = \sum_{i=1}^{m} \log(1 + e^{-y_i x_i^T w}) > 0$,

but $\lim_{n\to\infty} \sum_{i=1}^{m} \log(1 + e^{-y_i x_i^T nw_0}) = f(nw_0) = 0$

which means $f(w)$ must not have a global minimum.

(2) i) $f(w) = \sum_{i=1}^{m} \log(1 + e^{-y_i x_i^T w}) > \sum_{i=1, i\neq n}^{m} \log(1 + e^{-y_i x_i^T w}) + \log e^{-y_n x_n^T w}$

let $-y_n x_n^T w$ be the max $-y_i x_i^T w$

for the log part is non-negative, $f(w) > h(w)$ strictly

ii) let $w = (\cos\theta, \sin\theta)$

h is continuous and $h(\theta) = \max -y_i(x_{i1}\cos\theta + x_{i2}\sin\theta)$ is on a closed set $[0, 2\pi]$, so $h(\theta)$ must have a global minimum $h(\theta_0) = h(w_0)$

for any w, $\exists\, i$. $y_i x_i^T w < 0$

So $h(w) > 0$, $C = h(w_0) > 0$

iii) still consider $w \in S$, $f(nw) = nf(w)$, $n \in [0, +\infty)$

so we have $f(nw) \geq C$, $f(w) \geq C\|w\|$

iv) $f(w) \geq C\|w\|$, continuous

So $f(w) \to +\infty$ as $\|w\| \to +\infty$

$f(w)$ has a global minimum

(c) $$\nabla f(w) = \sum_{i=1}^{n} \frac{-y_i x_i^T e^{-y_i x_i^T w}}{1 + e^{-y_i x_i^T w}}$$

(d) whether it's linearly separable or not.

$\tilde{f}(w) \geq \frac{\lambda}{2}\|w\|_2^2$

So $f(w) \to +\infty$ as $\|w\| \to +\infty$

$f(w)$ has a global minimum