1. Write down the Wolfe dual of:

\[
\begin{align*}
\min_{x \in \mathbb{R}^n} & \quad x_1 - x_2 + x_2^2 \\
\text{s.t.} & \quad x \geq 0 \\
& \quad \frac{x_1^2}{4} + x_2^2 \leq 1
\end{align*}
\]

Reduce the number of variables of the dual problem. [Hint: Eliminate the vector of multipliers corresponding to \( x \geq 0 \).]

**Solution:** Recall that the Wolfe dual involves both the Lagrangian and the gradient of the Lagrangian, so we first write down the Lagrangian of the primal problem, and then find the gradient of the Lagrangian.

\[
L(x, \lambda) = x_1 - x_2 + x_2^2 - \lambda_1 x_1 - \lambda_2 x_2 + \lambda_3 \left( \frac{x_1^2}{4} + x_2^2 - 1 \right) \quad (1)
\]

\[
\nabla_{x_1} L(x, \lambda) = 1 - \lambda_1 + \lambda_3 \frac{x_1}{2} = 0 \quad (2)
\]

\[
\nabla_{x_2} L(x, \lambda) = -1 + 2x_2 - \lambda_2 + 2\lambda_3 x_2 = 0 \quad (3)
\]

From (2) we can get:

\[
\lambda_1 = 1 + \lambda_3 \frac{x_1}{2}. \quad (4)
\]

And from (3) we get:

\[
\lambda_2 = 2x_2 - 1 + 2\lambda_3 x_2. \quad (5)
\]

Substituting both (4) and (5) back into (1) and simplifying:

\[
L(x, \lambda_3) = -\lambda_3 \frac{x_1^2}{4} - (1 + \lambda_3)x_2^2 - \lambda_3 \quad (6)
\]

Since we constrain the Lagrange multipliers to be \( \geq 0 \) in the Wolfe dual, we know that \( \lambda_1 \geq 0 \) and \( \lambda_2 \geq 0 \). Therefore the Wolfe dual is:

\[
\max_{x, \lambda_3} \quad -\lambda_3 \frac{x_1^2}{4} - (1 + \lambda_3)x_2^2 - \lambda_3 \\
\text{s.t.} \quad \lambda_3 \geq 0 \\
& \quad 1 + \lambda_3 \frac{x_1}{2} \geq 0 \\
& \quad 2x_2 - 1 + 2\lambda_3 x_2 \geq 0
\]
2. Write down the Wolfe dual of:
\[
\min \quad f(x) \\
\text{s.t.} \quad x \geq 0 \\
Ax = b
\]
where \( f \) is smooth and \( A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m \). Simplify the dual problem by eliminating the vector of multipliers corresponding to \( x \geq 0 \), to obtain a nonlinear dual problem with linear constraints.

**Solution:** As with question (1), we first find the Lagrangian and its gradient.

\[
L(x, \lambda, \eta) = f(x) - \lambda^T x + \eta^T (b - Ax) \quad (7)
\]
\[
\nabla_x L(x, \lambda, \eta) = \nabla f(x) - \lambda - A^T \eta = 0 \quad (8)
\]

We can eliminate the vector \( \lambda \) by first re-arranging equation (8):

\[
\lambda = \nabla f(x) - A^T \eta, \quad (9)
\]

and then substituting (9) into the original Lagrangian equation:

\[
L(x, \eta) = f(x) - \nabla f(x)^T x + \eta^T b \quad (10)
\]

By observing that \( \lambda \geq 0 \), we get the Wolfe dual:

\[
\max \quad f(x) - \nabla f(x)^T x + \eta^T b \\
\text{s.t.} \quad \eta \geq 0 \\
A^T \eta \leq \nabla f(x)
\]