A

$$\text{prox}_{\ \gamma} f(x,x_0) = \text{arg min}_z \frac{1}{2\gamma}(z-x)^2 + z^T \nabla f(x_0) = \text{arg min}_z \frac{1}{2\gamma} z^2 - \frac{1}{\gamma} z^T x + z^T \nabla f(x_0) = x - \gamma \nabla f(x_0)$$

B

$$\begin{aligned}
\text{prox}_{\frac{1}{\gamma}} l(x) &= \text{arg min}_z \frac{\gamma}{2}(z-x)^2 + \frac{1}{2} z^T P z - q^T z \\
&= \text{arg min}_z \frac{1}{2} z^T (P + I\gamma) z - q^T z - \gamma x^T z \\
&= \text{arg min}_z \frac{1}{2}(z - (P+I\gamma)^{-1}(q+\gamma x))^T (P + I\lambda)(z - (P+I\gamma)^{-1}(q+\gamma x)) \\
&= (P+I\gamma)^{-1}(q+\gamma x)
\end{aligned}$$

$$\text{negative loglikelihood} = \frac{1}{2}(y-Ax)^T \Omega^{-1}(y-Ax) = \frac{1}{2} y^T \Omega^{-1} y - x^T A^T \Omega^{-1} y + \frac{1}{2} x^T A^T \Omega^{-1} A x$$

$$P = \Omega^{-1}$$

$$q = \Omega^{-1} A x$$

$$r = \frac{1}{2} x^T A^T \Omega^{-1} A x$$

C

$$\text{prox}_{\ \gamma} \phi(x) = \text{arg min}_z \frac{1}{2\gamma}(z-x)^2 + \tau|z| = \text{arg min}_z \frac{1}{2}(z-x)^2 + \tau\gamma|z| = S_{\tau\gamma}(x)$$

A

$$\begin{aligned}
& \text{arg min}_x l(x,x_0) + \phi(x) \\
= \;& \text{arg min}_x \frac{1}{2\gamma}(x - x_0)^2 + x \nabla l(x_0) + \phi(x) \\
= \;& \text{arg min}_x \frac{1}{2\gamma}(x - x_0 + \gamma \nabla l(x_0))^2 + \phi(x) \\
= \;& \text{prox}_\gamma \phi(x_0 - \gamma \nabla l(x_0))
\end{aligned}$$

B

$$\nabla_\beta \frac{1}{2n}(y - X\beta) = \frac{1}{n}(X^T X \beta + X^T y) := \Delta$$

$$\begin{aligned}
\beta_{t+1} &= \text{arg min}_{\beta_t} \frac{1}{2n}(y - X\beta_t)^2 + \lambda|\beta_t| \\
&= \text{prox}_\gamma \lambda|\beta_t - \gamma \Delta|_{\beta_t}| \\
&= S_{\lambda\gamma}(\beta_t - \gamma\Delta) \qquad (1)
\end{aligned}$$