Notation: $Y = (Y_1, Y_2, \dots, Y_n)^T \in \mathbb{R}^N$ and $U = (U_1, U_2, \dots, U_q)^T \in \mathbb{R}^q$ are random vectors (response and random/mixed effects); their realizations are denoted $y = (y_1, y_2, \dots, y_n)^T$ and $u = (u_1, u_2, \dots, u_q)^T$; $\beta \in \mathbb{R}^p$ is the vector of fixed effects; $X \in \mathbb{R}^{n\times p}$ and $Z \in \mathbb{R}^{n\times q}$ are model matrices, where x_i^T or z_i^T denote the ith row of X or Z respectively.

Generalized linear mixed model (GLMM):

$(Y_i|U=u) \sim \text{indep.} f_{Y_i|U}(y_i|u)$ for $i \in \{1, 2, \dots, n\}$, $U \sim f_U(u)$.
$f_{Y_i|U}(y_i|u) = h(y_i)\exp\left(\langle T(y_i), \eta\rangle - A(\eta)\right)$,
$\mu_i = \mathrm{E}(Y_i|U=u) = f(\eta)$,
$g(\mu_i) = x_i^T\beta + z_i^T u$.

Normality

$Z = 0$

Linear mixed model (LMM):

$(Y|U=u) \sim \mathcal{N}\left(X\beta + Zu, \sigma^2 I\right)$,
$U \sim \mathcal{N}(0, \Sigma(\theta))$.
Solution:
$\hat\beta_{\mathrm{ML}} = (X^T\hat V^{-1}X)^{-1}X^T\hat V^{-1}y$ is MLE, where
$\hat V = Z\Sigma_{\hat\theta_{\mathrm{ML}}}Z^T + \hat\sigma^2_{\mathrm{ML}}I$ is MLE of $\mathrm{Var}(Y)$.

Generalized linear model (GLM):

$Y_i \sim \text{indep.} f_{Y_i}(y_i)$ for $i \in \{1, 2, \dots, n\}$,
$f_{Y_i}(y_i) = h(y_i)\exp\left(\langle T(y_i), \eta\rangle - A(\eta)\right)$,
$\mu_i = \mathrm{E}(Y_i) = f(\eta)$,
$g(\mu_i) = x_i^T\beta$.

$Z = 0$

Normality

Linear model (LM):

$Y \sim \mathcal{N}\left(X\beta, \sigma^2 I\right)$.
Closed form solution:
$\hat\beta = (X^TX)^{-1}X^Ty$ is MLE,
UMVU and BLUE.

Normality, $\lambda = 0$

Normality, $\lambda = 0$

LASSO:

$\mathrm{E}(Y) = X\beta$, tuning parameter $\lambda \geq 0$,
$\hat\beta = \mathrm{argmin}_{b\in\mathbb{R}^p}\left\{\frac{1}{2n}\|y - Xb\|_2^2 + \lambda\|b\|_1\right\}$.

Ridge regression:

$\mathrm{E}(Y) = X\beta$, tuning parameter $\lambda \geq 0$,
$\hat\beta = \mathrm{argmin}_{b\in\mathbb{R}^p}\left\{\frac{1}{2n}\|y - Xb\|_2^2 + \lambda\|b\|_2^2\right\}$.
Closed form solution:
$\hat\beta = (X^TX + \lambda I)^{-1}X^Ty$.

$\alpha = 1$

$\alpha = 0$

Elastic net:

$\mathrm{E}(Y) = X\beta$, tuning parameters $\lambda \geq 0$ and $\alpha \in [0, 1]$,
$\hat\beta = \mathrm{argmin}_{b\in\mathbb{R}^p}\left\{\frac{1}{2}\|y - Xb\|_2^2 + \lambda\left[\frac{1}{2}(1-\alpha)\|b\|_2^2 + \alpha\|b\|_1\right]\right\}$.