Let us write the data model's density $p(x\eta)$ in exponential family form, and assume that the prior $p(\eta\phi)$ has an exponential family construction as well. align* p(x $\eta) = h(x) \exp\{\eta^T t(x) - a(\eta)\}$

Our goal is to find the form of a conjugate prior, given the form of the data model. To do this, we work with the posterior

align* p($\eta x, \phi) \propto p(x\eta) \quad p(\eta\phi) bayeslaw$

where in (1) we reorganize the likelihood to isolate terms in $\eta$, and then take $t(\eta)$ to have this same form in order to get conjugacy.

From this, we conclude enumerate

T he vector $t(\eta) = bmatrix \eta$

remark As this derivation shows, there are multiple possible conjugate priors, depending on the choice of $h$. For instance, whenever the normal distribution is conjugate, so is the truncated normal, since those distributions differ only in their carrier density (see Remark rk:truncated$_{n}ormal_{d}iffers_{f}rom_{n}ormal_{o}nly_{i}n_{t}erms_{o}f_{c}arrier_{d}ens$ $chosendominatingmeasure\nu$ on $\Phi$; for a discussion, see https://stats.stackexchange.com/questions/176668/can-anyone-explain-conjugate-priors-in-simplest-possible-terms$\underline{here}.rk : conjugate_{p}rior_{c}an_{h}ave_{a}ny_{d}esired_{c}arrier_{d}ensity$