**Wishart Distribution:** A $d \times d$ symmetric and positive-definite matrix $X$ has Wishart distribution with $\alpha$ degrees of freedom and parameter $\beta$, where $\alpha > (d - 1)/2$ and $\beta$ is a $d \times d$ symmetric and positive-definite matrix, its density is

$$f(x|\alpha, \beta) = \frac{|\beta|^{\alpha}}{\pi^{d(d-1)/4} \prod_{i=1}^{d} \Gamma(\alpha - \frac{i-1}{2})} |X|^{\alpha-(d+1)/2} \exp\{-\text{tr}(\beta X)\}$$

for positive-definite values of $X$ and 0 otherwise. $d = 1$ you get $G(\alpha, \beta)$ density. It can be shown that if $x_1, \cdots, x_n$ are a sample from $N_d(\mu, \Sigma)$ then $\sum_{i=1}^{n} (x_i - \mu)(x_i - \mu)^T \sim W_d(n/2, \Sigma^{-1}/2)$. So, an easy generation method for the $W_d(\alpha, \beta)$ is available from a sample of size $n = 2\alpha$ (integer) from the $N_d(\mu, \beta^{-1}/2)$.

**Multivariate Student’s t-distribution**

$X$ has $p$-variate Student’s t distribution with $\nu$ (a number) degrees of freedom, with mean $\mu$ ($\mu \in \mathbb{R}^p$) and scale parameter $\Sigma$ ($p \times p$ matrix) denoted by $t_\nu(\mu, \Sigma)$ with density

$$f(X|\nu, \mu, \Sigma) = \frac{\Gamma((\nu + p)/2)}{\Gamma(n/2)(n\pi)^{p/2}|\Sigma|^{-1/2}} \left[1 + \frac{1}{n}(X - \mu)^\prime \Sigma^{-1}(X - \mu)\right]^{-(\nu+p)/2}$$

$E(X) = \mu$, $\text{Var}(X) = n\nu\Sigma/(\nu - 2)$.

**Multivariate Model**

$y_1, y_2, \cdots, y_n$: sample from a $p$ dimensional multivariate normal distribution $N_p(\mu, \Sigma)$.

$$p(y_1, \cdots, y_n|\mu, \Sigma) \propto \prod_{i=1}^{n}|\Sigma|^{-1/2} \exp\{- (y_i - \mu)^\prime \Sigma^{-1} (y_i - \mu)/2 \}$$

$$= |\Sigma|^{-n/2} \exp\{- \sum_{i=1}^{n}(y_i - \mu)^\prime \Sigma^{-1} (y_i - \mu)/2 \}$$

Now expanding the quadratic forms we have

$$\sum_{i}(y_i - \mu)^\prime \Sigma^{-1}(y_i - \mu) = n\mu^\prime \Sigma^{-1}\mu - n\bar{y}^\prime \Sigma^{-1}\mu -$$

$$n\mu^\prime \Sigma^{-1}\bar{y} + \sum_{i}y_i^\prime \Sigma^{-1}y_i$$

$$= n(\mu - \bar{y})^\prime \Sigma^{-1}(\mu - \bar{y}) + r$$

where $\bar{y} = n^{-1} \sum_i y_i$ is the sample mean vector and $r = ntrbsigma^{-1}S$ where $S = n^{-1} \sum_i(y_i - \bar{y})(y_i - \bar{y})^\prime$ is the sample covariance matrix.
So the likelihood can be written as

\[
p(y_1, \cdots, y_n | \mu, \Sigma) \propto |\Sigma|^{-n/2} \exp\{-n(\mu - \bar{y})' \Sigma^{-1} (\mu - \bar{y})/2 - ntr\Sigma^{-1} S/2\}
\]

(1)

**Prior Distributions**

The prior distribution for $\Sigma$ must be a distribution over the space of all positive definite symmetric matrices. The conjugate family for multivariate normal likelihood is known as *Inverse-Wishart* distribution and as $\Sigma^{-1}$ has a Wishart distribution. $\Sigma \sim IW(A, d)$.

\[
p(\Sigma) \propto k^{-1} |A|^{d/2} |\Sigma|^{-(d+p+1)/2} \exp\{-tr\Sigma^{-1} A/2\}
\]

(2)

where $k = 2^{dp/2} \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma\{(d + 1 - i)/2\}$. $E(\Sigma) = (d - p + 1)^{-1} A$ provided $d > p + 1$ and also can find the variance-covariance elements of $\Sigma$. Siskind, V. (1972) Second moments of inverse Wishart-matrix elements. Biometrika, 59, 690-691.

\[
p(\mu, \Sigma) \propto |\Sigma|^{-(d+p+2)/2} \exp\{-c(\mu - a)' \Sigma^{-1}(\mu - a)/2 - (tr\Sigma^{-1} A)/2\}
\]

(3)

$p(\mu | \Sigma) \sim N_p(a, c^{-1} \Sigma)$. $p(\Sigma) \sim IW(A, d)$. This is known as Normal-inverse-Wishart distribution $NIW(A, d, a, c)$

\[
p(\mu, \Sigma | y_1 \cdots, y_n) \propto |\Sigma|^{-(d+p+2+n)/2} \exp(-Q/2)
\]

where $Q = tr\Sigma^{-1}(S + A) + c(\mu - a)' \Sigma^{-1}(\mu - a) + n(\mu - \bar{y})' \Sigma^{-1}(\mu - \bar{y})$.

\[
Q = tr\Sigma^{-1} A^* + c^*(\mu - a^*)' \Sigma^{-1}(\mu - a^*)
\]

where $a^* = (c + n)^{-1}(ca + n\bar{y})$, $c^* = c + n$, $A^* = A + S + cn(c + n)^{-1}(a - \bar{y})(a - \bar{y})'$. So the posterior is $NIW(A^*, d^*, a^*, c^*)$.

**Linear Model**

\[
y = X\beta + \epsilon
\]