Bayesian Inference

Loukia Meligkotsidou,
National and Kapodistrian University of Athens

MSc in Statistics and Operational Research,
Department of Mathematics

Outline of the course

This course provides theory and practice of the Bayesian approach to statistical inference. Applications are performed with the statistical package R.

Topics:

- Bayesian Updating through Bayes' Theorem
- Prior Distributions
- Multi-parameter Problems
- Summarizing Posterior Information
- The Multivariate Normal Model
- Prediction
- Asymptotics
- Markov chain Monte Carlo Methods

The Multivariate Normal Distribution

Definition

Let X be a $p \times 1$ real random vector. X follows the multivariate normal distribution with mean vector μ and covariance matrix Σ, i.e. $X \sim N_p(\mu, \Sigma)$, if and only if its pdf is

$$f(x \mid \mu, \Sigma) = (2\pi)^{-p/2}|\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right\},$$

where μ is a $p \times 1$ real vector and Σ is a $p \times p$ symmetric and positive definite matrix.

Random Sample from the Multivariate Normal Distribution

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ be an i.i.d sample from a random variable $\mathbf{X} \sim N_p(\mu, \Sigma)$.

The pdf of the ith observation, $\mathbf{x}_i$, is

$$f(\mathbf{x}_i \mid \mu, \Sigma) = (2\pi)^{-p/2} \mid \det(\Sigma) \mid^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x}_i - \mu)'\Sigma^{-1}(\mathbf{x}_i - \mu)\right],$$

where μ is the $p \times 1$ mean vector and Σ is the $p \times p$ symmetric and positive definite covariance matrix.

Random Sample from the Multivariate Normal Distribution

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ be an i.i.d sample from a random variable $\mathbf{X} \sim N_p(\mu, \Sigma)$.

The pdf of the ith observation, $\mathbf{x}_i$, is

$$f(\mathbf{x}_i \mid \mu, \Sigma) = (2\pi)^{-p/2} \mid \det(\Sigma) \mid^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x}_i - \mu)'\Sigma^{-1}(\mathbf{x}_i - \mu)\right],$$

where μ is the $p \times 1$ mean vector and Σ is the $p \times p$ symmetric and positive definite covariance matrix.

We will use the multivariate random sample to estimate the unknown parameters, μ, Σ, of the multivariate normal model.

The Likelihood Function

The likelihood function of the observed data is

$$\begin{aligned} L(\mu, \Sigma) &= \prod_{i=1}^{n} f(\mathbf{x}_i \mid \mu, \Sigma) \\ &= \prod_{i=1}^{n} (2\pi)^{-p/2} \mid \det(\Sigma) \mid^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x}_i - \mu)'\Sigma^{-1}(\mathbf{x}_i - \mu)\right] \\ &= (2\pi)^{-np/2} \mid \det(\Sigma) \mid^{-n/2} \\ &\quad \times \exp\left[-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \mu)'\Sigma^{-1}(\mathbf{x}_i - \mu)\right] \end{aligned}$$

Note that the likelihood function is well-defined only if $\det(\Sigma)$ is strictly positive. This reflects the assumption made above that Σ is positive definite, which implies that the search of a maximum likelihood estimator of Σ is restricted to the space of positive definite matrices.

The Log-likelihood Function

The log-likelihood function is

$$l(\mu, \Sigma) = -\frac{np}{2}\ln(2\pi) - \frac{n}{2}\ln(\det(\Sigma)) - \frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \mu)'\Sigma^{-1}(\mathbf{x}_i - \mu)$$

For convenience, the log-likelihood function can be also parameterised in terms of the precision matrix, Σ^{-1}, i.e.

$$l(\mu, \Sigma^{-1}) = -\frac{np}{2}\ln(2\pi) + \frac{n}{2}\ln(\det(\Sigma^{-1})) - \frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \mu)'\Sigma^{-1}(\mathbf{x}_i - \mu)$$

Preliminaries: Linear Algebra

- If c is a scalar, then the trace of c equals c: $tr(c) = c$.
- If two matrices A and B are such that the products AB and BA are both well defined, then $tr(AB) = tr(BA)$.
- The trace is a linear operator: if A and B are two matrices and a and b are two scalars, then $tr(aA + bB) = atr(A) + btr(B)$.
- The gradient of the trace of the product of two matrices A and B with respect to A is $\frac{\partial}{\partial A} tr(BA) = B'$.
- The gradient of the natural logarithm of the determinant of A is $\frac{\partial}{\partial A} \ln(\det(A)) = (A^{-1})'$.
- If $\mathbf{x}$ is a $p \times 1$ vector and A is a $p \times p$ symmetric matrix, then $\frac{\partial}{\partial \mathbf{x}} \mathbf{x}' A \mathbf{x} = 2A\mathbf{x}$.

The Log-likelihood Function Revisited

$$l(\mu, \Sigma^{-1}) = -\frac{np}{2}\ln(2\pi) + \frac{n}{2}\ln(\det(\Sigma^{-1})) - \frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \mu)'\Sigma^{-1}(\mathbf{x}_i - \mu)$$

The Log-likelihood Function Revisited

$$l(\mu, \Sigma^{-1}) = -\frac{np}{2}\ln(2\pi) + \frac{n}{2}\ln(\det(\Sigma^{-1})) - \frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \mu)'\Sigma^{-1}(\mathbf{x}_i - \mu)$$

$$= \frac{np}{2}\ln(2\pi) + \frac{n}{2}\ln(\det(\Sigma^{-1})) - \frac{1}{2}tr\left[\sum_{i=1}^{n}(\mathbf{x}_i - \mu)'\Sigma^{-1}(\mathbf{x}_i - \mu)\right]$$

The Log-likelihood Function Revisited

$$l(\mu, \Sigma^{-1}) = -\frac{np}{2}\ln(2\pi) + \frac{n}{2}\ln(\det(\Sigma^{-1})) - \frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \mu)'\Sigma^{-1}(\mathbf{x}_i - \mu)$$

$$= \frac{np}{2}\ln(2\pi) + \frac{n}{2}\ln(\det(\Sigma^{-1})) - \frac{1}{2}tr\left[\sum_{i=1}^{n}(\mathbf{x}_i - \mu)'\Sigma^{-1}(\mathbf{x}_i - \mu)\right]$$

$$= \frac{np}{2}\ln(2\pi) + \frac{n}{2}\ln(\det(\Sigma^{-1})) - \frac{1}{2}tr\left[\sum_{i=1}^{n}(\mathbf{x}_i - \mu)(\mathbf{x}_i - \mu)'\Sigma^{-1}\right]$$

The Log-likelihood Function Revisited

$$l(\mu, \Sigma^{-1}) = -\frac{np}{2}\ln(2\pi) + \frac{n}{2}\ln(\det(\Sigma^{-1})) - \frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \mu)'\Sigma^{-1}(\mathbf{x}_i - \mu)$$

$$= \frac{np}{2}\ln(2\pi) + \frac{n}{2}\ln(\det(\Sigma^{-1})) - \frac{1}{2}tr\left[\sum_{i=1}^{n}(\mathbf{x}_i - \mu)'\Sigma^{-1}(\mathbf{x}_i - \mu)\right]$$

$$= \frac{np}{2}\ln(2\pi) + \frac{n}{2}\ln(\det(\Sigma^{-1})) - \frac{1}{2}tr\left[\sum_{i=1}^{n}(\mathbf{x}_i - \mu)(\mathbf{x}_i - \mu)'\Sigma^{-1}\right]$$

$$= \frac{np}{2}\ln(2\pi) + \frac{n}{2}\ln(\det(\Sigma^{-1})) - \frac{1}{2}\sum_{i=1}^{n} tr\left[(\mathbf{x}_i - \mu)(\mathbf{x}_i - \mu)'\Sigma^{-1}\right]$$

Maximum Likelihood Estimation

To obtain the maximum likelihood estimators of μ, Σ we need to solve the system

$$\frac{\partial l(\mu, \Sigma^{-1})}{\partial \mu} \Big|_{\hat{\mu}, \hat{\Sigma}^{-1}} = \mathbf{0} \quad (1)$$

$$\frac{\partial l(\mu, \Sigma^{-1})}{\partial \Sigma^{-1}} \Big|_{\hat{\mu}, \hat{\Sigma}^{-1}} = \mathbf{0} \quad (2)$$

Maximum Likelihood Estimation

To obtain the maximum likelihood estimators of μ, Σ we need to solve the system

$$\begin{aligned} \frac{\partial l(\mu, \Sigma^{-1})}{\partial \mu} \Big|_{\hat{\mu}, \hat{\Sigma}^{-1}} &= \mathbf{0} \quad (1) \\ \frac{\partial l(\mu, \Sigma^{-1})}{\partial \Sigma^{-1}} \Big|_{\hat{\mu}, \hat{\Sigma}^{-1}} &= \mathbf{0} \quad (2) \end{aligned}$$

$$(1) \Rightarrow -\hat{\Sigma}^{-1} \sum_{i=1}^{n} (\mathbf{x}_i - \hat{\mu}) = \mathbf{0} \Rightarrow \hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i = \bar{\mathbf{x}}$$

Maximum Likelihood Estimation

To obtain the maximum likelihood estimators of μ, Σ we need to solve the system

$$\begin{aligned} \frac{\partial l(\mu, \Sigma^{-1})}{\partial \mu} \Big|_{\hat{\mu}, \hat{\Sigma}^{-1}} &= \mathbf{0} \quad (1) \\ \frac{\partial l(\mu, \Sigma^{-1})}{\partial \Sigma^{-1}} \Big|_{\hat{\mu}, \hat{\Sigma}^{-1}} &= \mathbf{0} \quad (2) \end{aligned}$$

$$(1) \Rightarrow -\hat{\Sigma}^{-1} \sum_{i=1}^{n} (\mathbf{x}_i - \hat{\mu}) = \mathbf{0} \Rightarrow \hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i = \bar{\mathbf{x}}$$

$$(2) \Rightarrow \frac{n}{2}\hat{\Sigma} - \frac{1}{2} \sum_{i=1}^{n} (\mathbf{x}_i - \hat{\mu})(\mathbf{x}_i - \hat{\mu})' \Rightarrow \hat{\Sigma} = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'$$

Conjugate Analysis for the Multivariate Normal Model

The Inverse Wishart Distribution

Let Σ be a $p \times p$ symmetric and positive definite matrix. Σ follows the inverse Wishart distribution, i.e. $\Sigma \sim IW(A, d)$, if and only if its pdf is

$$f(\Sigma) = k^{-1}|A|^{d/2}|\Sigma|^{-(d+p+1)/2} \exp\left\{-\frac{1}{2}tr\Sigma^{-1}A\right\},$$

where $k = 2^{dp/2}\pi^{p(p-1)/4}\prod_{i=1}^{p}\Gamma((d+1-i)/2)$, A is a $p \times p$ symmetric and positive definite matrix and d is a positive scalar.

Conjugate Analysis for the Multivariate Normal Model

The Normal - Inverse Wishart Distribution

$(\mu, \Sigma) \sim NIW(A, d, a, c)$, where A is a $p \times p$ symmetric and positive definite matrix, a is a $p \times 1$ real vector and d, c are positive scalars.

Joint distribution for (μ, Σ) of the form

$$f(\mu, \Sigma) = f(\mu \mid \Sigma) f(\Sigma)$$

where $f(\Sigma) \equiv IW(A, d)$ and $f(\mu \mid \Sigma) \equiv N_p(a, c^{-1}\Sigma)$.

Conjugate Analysis for the Multivariate Normal Model

Let $X_1, X_2, \ldots, X_n$ be a random sample from the multivariate normal distribution $N_p(\mu, \Sigma)$.

Likelihood of $x = (x_1, x_2, \ldots, x_n)$:

$$
\begin{aligned}
f(x \mid \mu, \Sigma) &\propto \prod_{i=1}^{n} |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(x_i - \mu)'\Sigma^{-1}(x_i - \mu)\right\} \\
&= |\Sigma|^{-n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)'\Sigma^{-1}(x_i - \mu)\right\}
\end{aligned}
$$

Conjugate Analysis for the Multivariate Normal Model

Let $X_1, X_2, \ldots, X_n$ be a random sample from the multivariate normal distribution $N_p(\mu, \Sigma)$.

Likelihood of $x = (x_1, x_2, \ldots, x_n)$:

$$
\begin{aligned}
f(x \mid \mu, \Sigma) &\propto \prod_{i=1}^{n} |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(x_i - \mu)'\Sigma^{-1}(x_i - \mu)\right\} \\
&= |\Sigma|^{-n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)'\Sigma^{-1}(x_i - \mu)\right\} \\
&= |\Sigma|^{-n/2} \exp\left\{-\frac{n}{2}(\mu - \bar{x})'\Sigma^{-1}(\mu - \bar{x}) - \frac{n}{2} tr\Sigma^{-1}S, \right\}
\end{aligned}
$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n}$, $S = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})'$.

Conjugate Analysis for the Multivariate Normal Model

Let $\sum_{i=1}^{n}(x_i - \mu)'\Sigma^{-1}(x_i - \mu) = C$. Then

$$C = n\mu'\Sigma^{-1}\mu - n\bar{x}'\Sigma^{-1}\mu - n\mu'\Sigma^{-1}\bar{x} + \sum_{i=1}^{n} x_i'\Sigma^{-1}x_i$$

Conjugate Analysis for the Multivariate Normal Model

Let $\sum_{i=1}^{n}(x_i - \mu)'\Sigma^{-1}(x_i - \mu) = C$. Then

$$\begin{aligned} C &= n\mu'\Sigma^{-1}\mu - n\bar{x}'\Sigma^{-1}\mu - n\mu'\Sigma^{-1}\bar{x} + \sum_{i=1}^{n} x_i'\Sigma^{-1}x_i \\ &= n(\mu - \bar{x})'\Sigma^{-1}(\mu - \bar{x}) - n\bar{x}'\Sigma^{-1}\bar{x} + \sum_{i=1}^{n} x_i'\Sigma^{-1}x_i \end{aligned}$$

Conjugate Analysis for the Multivariate Normal Model

Let $\sum_{i=1}^{n}(x_i - \mu)'\Sigma^{-1}(x_i - \mu) = C$. Then

$$\begin{aligned} C &= n\mu'\Sigma^{-1}\mu - n\bar{x}'\Sigma^{-1}\mu - n\mu'\Sigma^{-1}\bar{x} + \sum_{i=1}^{n} x_i'\Sigma^{-1}x_i \\ &= n(\mu - \bar{x})'\Sigma^{-1}(\mu - \bar{x}) - n\bar{x}'\Sigma^{-1}\bar{x} + \sum_{i=1}^{n} x_i'\Sigma^{-1}x_i \\ &= n(\mu - \bar{x})'\Sigma^{-1}(\mu - \bar{x}) + \sum_{i=1}^{n}(x_i - \bar{x})'\Sigma^{-1}(x_i - \bar{x}) \end{aligned}$$

Conjugate Analysis for the Multivariate Normal Model

Let $\sum_{i=1}^{n}(x_i - \mu)'\Sigma^{-1}(x_i - \mu) = C$. Then

$$
\begin{aligned}
C &= n\mu'\Sigma^{-1}\mu - n\bar{x}'\Sigma^{-1}\mu - n\mu'\Sigma^{-1}\bar{x} + \sum_{i=1}^{n} x_i'\Sigma^{-1}x_i \\
&= n(\mu - \bar{x})'\Sigma^{-1}(\mu - \bar{x}) - n\bar{x}'\Sigma^{-1}\bar{x} + \sum_{i=1}^{n} x_i'\Sigma^{-1}x_i \\
&= n(\mu - \bar{x})'\Sigma^{-1}(\mu - \bar{x}) + \sum_{i=1}^{n}(x_i - \bar{x})'\Sigma^{-1}(x_i - \bar{x}) \\
&= n(\mu - \bar{x})'\Sigma^{-1}(\mu - \bar{x}) + \sum_{i=1}^{n} tr\Sigma^{-1}(x_i - \bar{x})(x_i - \bar{x})'
\end{aligned}
$$

Conjugate Analysis for the Multivariate Normal Model

Let $\sum_{i=1}^{n}(x_i-\mu)'\Sigma^{-1}(x_i-\mu)=C$. Then

$$\begin{aligned}
C &= n\mu'\Sigma^{-1}\mu - n\bar{x}'\Sigma^{-1}\mu - n\mu'\Sigma^{-1}\bar{x} + \sum_{i=1}^{n} x_i'\Sigma^{-1}x_i \\
&= n(\mu-\bar{x})'\Sigma^{-1}(\mu-\bar{x}) - n\bar{x}'\Sigma^{-1}\bar{x} + \sum_{i=1}^{n} x_i'\Sigma^{-1}x_i \\
&= n(\mu-\bar{x})'\Sigma^{-1}(\mu-\bar{x}) + \sum_{i=1}^{n}(x_i-\bar{x})'\Sigma^{-1}(x_i-\bar{x}) \\
&= n(\mu-\bar{x})'\Sigma^{-1}(\mu-\bar{x}) + \sum_{i=1}^{n} tr\Sigma^{-1}(x_i-\bar{x})(x_i-\bar{x})' \\
&= n(\mu-\bar{x})'\Sigma^{-1}(\mu-\bar{x}) + ntr\Sigma^{-1}S
\end{aligned}$$

where $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}$, $S=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})'$.

Conjugate Analysis for the Multivariate Normal Model

Likelihood of x:

$$f(x \mid \mu, \Sigma) \propto |\Sigma|^{-n/2} \exp\left\{-\frac{n}{2}(\mu - \bar{x})'\Sigma^{-1}(\mu - \bar{x}) - \frac{n}{2} tr\Sigma^{-1}S\right\}$$

Conjugate prior for (μ, Σ): Normal-Inverse Wishart

$$\begin{aligned} f(\mu, \Sigma) &\propto |\Sigma|^{-(d+p+2)/2} \exp\left\{-\frac{c}{2}(\mu - a)'\Sigma^{-1}(\mu - a) - \frac{1}{2} tr\Sigma^{-1}A\right\} \\ &\equiv NIW(A, d, a, c), \end{aligned}$$

where A is a $p \times p$ symmetric and positive definite matrix, a is a $p \times 1$ real vector and d, c are positive scalars.

The Joint Posterior

$$f(\mu, \Sigma \mid x) \propto f(x \mid \mu, \Sigma) f(\mu, \Sigma) \propto |\Sigma|^{-(d+p+2+n)/2} \exp\{-Q/2\},$$

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

The Joint Posterior

$$f(\mu, \Sigma \mid x) \propto f(x \mid \mu, \Sigma) f(\mu, \Sigma) \propto |\Sigma|^{-(d+p+2+n)/2} \exp\{-Q/2\},$$

$$\begin{aligned} Q &= tr\Sigma^{-1}(S+A) + c(\mu - a)'\Sigma^{-1}(\mu - a) + n(\mu - \bar{x})'\Sigma^{-1}(\mu - \bar{x}) \\ &= tr\Sigma^{-1}(S+A) + c\mu'\Sigma^{-1}\mu + ca'\Sigma^{-1}a - c\mu'\Sigma^{-1}a - ca'\Sigma^{-1}\mu \\ &\quad + n\mu'\Sigma^{-1}\mu + n\bar{x}'\Sigma^{-1}\bar{x} - n\mu'\Sigma^{-1}\bar{x} - n\bar{x}'\Sigma^{-1}\mu \end{aligned}$$

The Joint Posterior

$$f(\mu, \Sigma \mid x) \propto f(x \mid \mu, \Sigma) f(\mu, \Sigma) \propto |\Sigma|^{-(d+p+2+n)/2} \exp\{-Q/2\},$$

$$\begin{aligned}
Q &= tr\Sigma^{-1}(S+A) + c(\mu - a)'\Sigma^{-1}(\mu - a) + n(\mu - \bar{x})'\Sigma^{-1}(\mu - \bar{x}) \\
&= tr\Sigma^{-1}(S+A) + c\mu'\Sigma^{-1}\mu + ca'\Sigma^{-1}a - c\mu'\Sigma^{-1}a - ca'\Sigma^{-1}\mu \\
&\quad + n\mu'\Sigma^{-1}\mu + n\bar{x}'\Sigma^{-1}\bar{x} - n\mu'\Sigma^{-1}\bar{x} - n\bar{x}'\Sigma^{-1}\mu \\
&= tr\Sigma^{-1}(S+A) + (c+n)\mu'\Sigma^{-1}\mu - (ca + n\bar{x})'\Sigma^{-1}\mu \\
&\quad - \mu'\Sigma^{-1}(ca + n\bar{x}) + ca'\Sigma^{-1}a + n\bar{x}'\Sigma^{-1}\bar{x}
\end{aligned}$$

The Joint Posterior

$$f(\mu, \Sigma \mid x) \propto f(x \mid \mu, \Sigma) f(\mu, \Sigma) \propto |\Sigma|^{-(d+p+2+n)/2} \exp\{-Q/2\},$$

$$\begin{aligned}
Q &= tr\Sigma^{-1}(S+A) + c(\mu - a)'\Sigma^{-1}(\mu - a) + n(\mu - \bar{x})'\Sigma^{-1}(\mu - \bar{x}) \\
&= tr\Sigma^{-1}(S+A) + c\mu'\Sigma^{-1}\mu + ca'\Sigma^{-1}a - c\mu'\Sigma^{-1}a - ca'\Sigma^{-1}\mu \\
&\quad + n\mu'\Sigma^{-1}\mu + n\bar{x}'\Sigma^{-1}\bar{x} - n\mu'\Sigma^{-1}\bar{x} - n\bar{x}'\Sigma^{-1}\mu \\
&= tr\Sigma^{-1}(S+A) + (c+n)\mu'\Sigma^{-1}\mu - (ca + n\bar{x})'\Sigma^{-1}\mu \\
&\quad - \mu'\Sigma^{-1}(ca + n\bar{x}) + ca'\Sigma^{-1}a + n\bar{x}'\Sigma^{-1}\bar{x} \\
&= \ldots \\
&= tr\Sigma^{-1}(A^*) + c^*(\mu - a^*)'\Sigma^{-1}(\mu - a^*),
\end{aligned}$$

where $c^* = c + n$, $a^* = (c+n)^{-1}(ca + n\bar{x})$ and $A^* = A + S + nc(c+n)^{-1}(a - \bar{x})(a - \bar{x})'$.

Joint and Conditional Posteriors

$$f(\mu, \Sigma \mid x) \propto |\Sigma|^{-\frac{d^*+p+2}{2}} \exp\{-\frac{c^*}{2}(\mu - a^*)'\Sigma^{-1}(\mu - a^*) - \frac{1}{2} tr\Sigma^{-1}A^*\}$$
$$\equiv NIW(A^*, d^*, a^*, c^*),$$

where $d^* = d + n$, $c^* = c + n$, $a^* = (c + n)^{-1}(ca + n\bar{x})$ and $A^* = A + S + nc(c + n)^{-1}(a - \bar{x})(a - \bar{x})'$.

Joint and Conditional Posteriors

$$\begin{aligned} f(\mu, \Sigma \mid x) &\propto |\Sigma|^{-\frac{d^*+p+2}{2}} \exp\{-\frac{c^*}{2}(\mu - a^*)'\Sigma^{-1}(\mu - a^*) - \frac{1}{2} tr\Sigma^{-1}A^*\} \\ &\equiv NIW(A^*, d^*, a^*, c^*), \end{aligned}$$

where $d^* = d + n$, $c^* = c + n$, $a^* = (c + n)^{-1}(ca + n\bar{x})$ and $A^* = A + S + nc(c + n)^{-1}(a - \bar{x})(a - \bar{x})'$.

$$f(\mu \mid \Sigma, x) \propto \exp\{-\frac{c^*}{2}(\mu - a^*)'\Sigma^{-1}(\mu - a^*)\} \equiv N_p(a^*, c\Sigma^{-1})$$

Joint and Conditional Posteriors

$$f(\mu, \Sigma \mid x) \propto |\Sigma|^{-\frac{d^*+p+2}{2}} \exp\{-\frac{c^*}{2}(\mu - a^*)'\Sigma^{-1}(\mu - a^*) - \frac{1}{2} tr\Sigma^{-1}A^*\}$$
$$\equiv NIW(A^*, d^*, a^*, c^*),$$

where $d^* = d + n$, $c^* = c + n$, $a^* = (c + n)^{-1}(ca + n\bar{x})$ and $A^* = A + S + nc(c + n)^{-1}(a - \bar{x})(a - \bar{x})'$.

$$f(\mu \mid \Sigma, x) \propto \exp\{-\frac{c^*}{2}(\mu - a^*)'\Sigma^{-1}(\mu - a^*)\} \equiv N_p(a^*, c\Sigma^{-1})$$

$$f(\Sigma \mid \mu, x) \propto |\Sigma|^{-\frac{d^*+p+2}{2}} \exp\{-\frac{c^*}{2}(\mu - a^*)'\Sigma^{-1}(\mu - a^*) - \frac{1}{2} tr\Sigma^{-1}A^*\}$$
$$= |\Sigma|^{-\frac{d^*+p+2}{2}} \exp\{-\frac{1}{2} tr\Sigma^{-1}A^*_\mu\} \equiv IW(A^*_\mu, d^* + 1),$$

where $A^*_\mu = A^* + c^*(\mu - a^*)(\mu - a^*)'$.

Marginal Posterior of Σ

$$
\begin{aligned}
f(\Sigma|x) &\propto \int f(\mu, \Sigma \mid x) d\mu \\
&\propto \int |\Sigma|^{-\frac{d^*+p+2}{2}} \exp\{-\frac{c^*}{2}(\mu - a^*)'\Sigma^{-1}(\mu - a^*) - \frac{1}{2} tr\Sigma^{-1}A^*\}d
\end{aligned}
$$

Marginal Posterior of Σ

$$
\begin{aligned}
f(\Sigma|x) &\propto \int f(\mu,\Sigma \mid x)d\mu \\
&\propto \int |\Sigma|^{-\frac{d^*+p+2}{2}} \exp\{-\frac{c^*}{2}(\mu-a^*)'\Sigma^{-1}(\mu-a^*) - \frac{1}{2}tr\Sigma^{-1}A^*\}d \\
&= |\Sigma|^{-\frac{d^*+p+2}{2}} \exp\{-\frac{1}{2}tr\Sigma^{-1}A^*\} \\
&\quad \times \int \exp\{-\frac{c^*}{2}(\mu-a^*)'\Sigma^{-1}(\mu-a^*)\}d\mu
\end{aligned}
$$

Marginal Posterior of Σ

$$
\begin{aligned}
f(\Sigma|x) &\propto \int f(\mu,\Sigma \mid x)d\mu \\
&\propto \int |\Sigma|^{-\frac{d^*+p+2}{2}} \exp\{-\frac{c^*}{2}(\mu-a^*)'\Sigma^{-1}(\mu-a^*) - \frac{1}{2}tr\Sigma^{-1}A^*\}d \\
&= |\Sigma|^{-\frac{d^*+p+2}{2}} \exp\{-\frac{1}{2}tr\Sigma^{-1}A^*\} \\
&\quad \times \int \exp\{-\frac{c^*}{2}(\mu-a^*)'\Sigma^{-1}(\mu-a^*)\}d\mu \\
&\propto |\Sigma|^{-\frac{d^*+p+2}{2}} \exp\{-\frac{1}{2}tr\Sigma^{-1}A^*\} \times |\Sigma|^{1/2} \\
&= |\Sigma|^{-\frac{d^*+p+1}{2}} \exp\{-\frac{1}{2}tr\Sigma^{-1}A^*\} \equiv IW(A^*,d^*)
\end{aligned}
$$

Marginal Posterior of μ

$$
\begin{aligned}
f(\mu|x) &\propto \int f(\mu, \Sigma \mid x) d\Sigma \\
&\propto \int |\Sigma|^{-\frac{d^*+p+2}{2}} \exp\{-\frac{1}{2} tr\Sigma^{-1} A_{\mu}^{*}\} d\Sigma
\end{aligned}
$$

Marginal Posterior of μ

$$
\begin{aligned}
f(\mu|x) &\propto \int f(\mu, \Sigma \mid x) d\Sigma \\
&\propto \int |\Sigma|^{-\frac{d^*+p+2}{2}} \exp\{-\frac{1}{2} tr\Sigma^{-1} A_\mu^*\} d\Sigma \\
&\propto |A_\mu^*|^{(d^*+1)/2} \\
&= |A^* + c^*(\mu - a^*)(\mu - a^*)'|^{-(d^*+1)/2}
\end{aligned}
$$

Marginal Posterior of μ

$$
\begin{aligned}
f(\mu|x) &\propto \int f(\mu, \Sigma \mid x) d\Sigma \\
&\propto \int |\Sigma|^{-\frac{d^*+p+2}{2}} \exp\{-\frac{1}{2} tr \Sigma^{-1} A_\mu^*\} d\Sigma \\
&\propto |A_\mu^*|^{(d^*+1)/2} \\
&= |A^* + c^*(\mu - a^*)(\mu - a^*)'|^{-(d^*+1)/2} \\
&\equiv t_{d^*+1-p}(a^*, \frac{1}{c^*} A^*)
\end{aligned}
$$

Marginal Posterior of μ

$$
\begin{aligned}
f(\mu|x) &\propto \int f(\mu, \Sigma \mid x) d\Sigma \\
&\propto \int |\Sigma|^{-\frac{d^*+p+2}{2}} \exp\{-\frac{1}{2} tr\Sigma^{-1} A^*_\mu\} d\Sigma \\
&\propto |A^*_\mu|^{(d^*+1)/2} \\
&= |A^* + c^*(\mu - a^*)(\mu - a^*)'|^{-(d^*+1)/2} \\
&\equiv t_{d^*+1-p}(a^*, \frac{1}{c^*} A^*)
\end{aligned}
$$

This is a non-central multivariate Student-t distribution with $d^* + 1 - p$ degrees of freedom, mean equal to a^* and covariance matrix equal to $\frac{1}{(d^*-p-1)c^*} A^*$.