Derivation of the ML discrepancy function from likelihoods[1]

Kristopher J. Preacher
Vanderbilt University

This document grew out of an answer to a student question in my Factor Analysis course. I thought it would be a handy reference for others. The question was: "How do you get from the likelihood formula to a discrepancy function?" I could not find the details of the calculations in any reference, so derivations are provided here.

There are different answers to the question, depending on whether we want to include structured means in the model and on what sort of likelihood we use. CEFA (Browne, Cudeck, Tateneni, & Mels, 2008) uses a discrepancy function based on the Wishart distribution—sort of like a multivariate χ^2 distribution describing sampling properties of covariance matrices. This kind of likelihood is called *maximum Wishart likelihood* (MWL). Other situations call for using a likelihood based on the joint multivariate normal distribution (MVN). For example, FIML estimation in SEM (e.g., in Mplus; Muthén, 2004) uses the MVN likelihood formulation because it is based on casewise sample data rather than on sample covariance matrices, which require complete (or imputed) data. I will address MVN likelihood first, both without and with a mean structure.

Throughout, we make use of the fact that the maximum likelihood estimate of $\mathbf{\Sigma}$ is $\mathbf{W} = N^{-1}\sum(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'$, which involves dividing by N rather than the more familiar $N-1$. The covariance matrix $\mathbf{W}$ is used in MVN discrepancy functions despite being a little biased because it has smaller mean squared error. The traditional sample covariance matrix, on the other hand, involves dividing by $N-1$: $\mathbf{S} = (N-1)^{-1}\sum(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'$. The covariance matrix $\mathbf{S}$ is used in MWL discrepancy functions.

Discrepancy based on the multivariate normal distribution

First, define two likelihoods, one for the *proposed model* (Model *M*) and one for a more general comparison model. The comparison model here will be the *saturated model* (Model 0), one that completely accounts for the observed data (i.e., yields perfect fit).

The multivariate normal joint likelihood function for the **proposed model** is:

$$L_0 = \prod_{i=1}^{N} \frac{|\mathbf{\Sigma}_0|^{-\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}} \exp\left[-\frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu}_0)' \mathbf{\Sigma}_0^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_0)\right]$$

[1] My thanks to Guangjian Zhang and Patrick O'Keefe for providing careful readings and catching some errors.

The MVN joint likelihood function for the **saturated model** is:

$$L_1 = \prod_{i=1}^{N} \frac{|\mathbf{\Sigma}_1|^{-\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}} \exp\left[-\frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu}_1)' \mathbf{\Sigma}_1^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_1)\right]$$

Define the **likelihood ratio** as:

$$LR = \frac{L_0}{L_1}$$

LR is always between 0 and 1, so $-2\ln LR$ will always be positive.

Define the **discrepancy function** $F_{\text{ML}}(\mathbf{W}, \mathbf{\Sigma}_0)$ using:

$$N \times F_{\text{ML}}(\mathbf{W}, \mathbf{\Sigma}_0) = -2\ln LR$$
$$N \times F_{\text{ML}}(\mathbf{W}, \mathbf{\Sigma}_0) = -2\ln \frac{L_0}{L_1}$$
$$N \times F_{\text{ML}}(\mathbf{W}, \mathbf{\Sigma}_0) = -2\ln L_0 + 2\ln L_1$$

We need to know what the quantities $-2\ln L_0$ and $-2\ln L_1$ equal. Comments in blue indicate brief explanations of non-obvious algebraic operations. First, for the proposed model...

$$-2\ln L_0 = -2\ln\left\{\prod_{i=1}^{N} \frac{|\mathbf{\Sigma}_0|^{-\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}} \exp\left[-\frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu}_0)' \mathbf{\Sigma}_0^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_0)\right]\right\}$$

$$-2\ln L_0 = -2\sum_{i=1}^{N} \ln\left\{\frac{|\mathbf{\Sigma}_0|^{-\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}} \exp\left[-\frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu}_0)' \mathbf{\Sigma}_0^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_0)\right]\right\}$$

$$-2\ln L_0 = -2\sum_{i=1}^{N} \left\{\ln\left\{\frac{|\mathbf{\Sigma}_0|^{-\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}}\right\} + \ln\left\{\exp\left[-\frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu}_0)' \mathbf{\Sigma}_0^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_0)\right]\right\}\right\}$$

$$-2\ln L_0 = -2\sum_{i=1}^{N} \left\{\ln\left\{|\mathbf{\Sigma}_0|^{-\frac{1}{2}}\right\} - \ln\left\{(2\pi)^{\frac{p}{2}}\right\} - \frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu}_0)' \mathbf{\Sigma}_0^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_0)\right\}$$

$$-2\ln L_0 = -2\sum_{i=1}^{N}\left\{-\frac{1}{2}\ln\left|\boldsymbol{\Sigma}_0\right| - \frac{p}{2}\ln(2\pi) - \frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_0)\right\}$$

$$-2\ln L_0 = \sum_{i=1}^{N}\ln\left|\boldsymbol{\Sigma}_0\right| + \sum_{i=1}^{N} p\ln(2\pi) + \sum_{i=1}^{N}(\mathbf{x}_i - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_0)$$

$$-2\ln L_0 = N\ln\left|\boldsymbol{\Sigma}_0\right| + Np\ln(2\pi) + \sum_{i=1}^{N}(\mathbf{x}_i - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_0)$$

$$-2\ln L_0 = N\ln\left|\boldsymbol{\Sigma}_0\right| + Np\ln(2\pi) + \sum_{i=1}^{N} tr\left\{(\mathbf{x}_i - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_0)\right\} \text{ b/c scalar = tr(scalar)}$$

$$-2\ln L_0 = N\ln\left|\boldsymbol{\Sigma}_0\right| + Np\ln(2\pi) + \sum_{i=1}^{N} tr\left\{(\mathbf{x}_i - \boldsymbol{\mu}_0)(\mathbf{x}_i - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}\right\} \text{ identity}$$

$$-2\ln L_0 = N\ln\left|\boldsymbol{\Sigma}_0\right| + Np\ln(2\pi) + N\frac{1}{N}\sum_{i=1}^{N} tr\left\{(\mathbf{x}_i - \boldsymbol{\mu}_0)(\mathbf{x}_i - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}\right\} \text{ adding } N\text{s strategically}$$

$$-2\ln L_0 = N\ln\left|\boldsymbol{\Sigma}_0\right| + Np\ln(2\pi) + Ntr\left\{\boldsymbol{\Sigma}_0^{-1}\frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i - \boldsymbol{\mu}_0)(\mathbf{x}_i - \boldsymbol{\mu}_0)'\right\} \text{ remove constant } \boldsymbol{\Sigma}^{-1}\text{; tr(sum) = sum(tr)}$$

$$-2\ln L_0 = N\ln\left|\boldsymbol{\Sigma}_0\right| + Np\ln(2\pi) + Ntr\left(\boldsymbol{\Sigma}_0^{-1}\mathbf{W}\right) \text{ definition of a covariance matrix}$$

In the last step, we implicitly assumed that the mean structure is uninteresting to us, and thus that $\boldsymbol{\mu}_0 = \overline{\mathbf{x}}$, allowing us to turn $\frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i - \boldsymbol{\mu}_0)(\mathbf{x}_i - \boldsymbol{\mu}_0)'$ (which includes parameters) into a *sample* covariance matrix $\mathbf{W}$ involving no parameters.

Next, for the saturated model...

$$-2\ln L_1 = -2\ln\left\{\prod_{i=1}^{N}\frac{\left|\boldsymbol{\Sigma}_1\right|^{-\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}}\exp\left[-\frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu}_1)'\boldsymbol{\Sigma}_1^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_1)\right]\right\}$$

(repeat a few steps from the previous derivation)

$$-2\ln L_1 = N\ln\left|\boldsymbol{\Sigma}_1\right| + Np\ln(2\pi) + Ntr\left\{\boldsymbol{\Sigma}_1^{-1}\frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i - \boldsymbol{\mu}_1)(\mathbf{x}_i - \boldsymbol{\mu}_1)'\right\}$$

$$-2\ln L_1 = N\ln\left|\boldsymbol{\Sigma}_1\right| + Np\ln(2\pi) + Ntr\left(\boldsymbol{\Sigma}_1^{-1}\mathbf{W}\right) \text{ definition of a covariance matrix}$$

$$-2\ln L_1 = N\ln\left|\mathbf{W}\right| + Np\ln(2\pi) + Ntr\left(\mathbf{W}^{-1}\mathbf{W}\right) \text{ note: } \mathbf{W} = \boldsymbol{\Sigma} \text{ for the saturated model}$$

Again, we assumed that $\boldsymbol{\mu}_1 = \overline{\mathbf{x}}$. This will be true for the saturated model regardless of whether or not we are concerned with the mean structure.

Now we are in a position to compute $F_{\text{ML}}(\mathbf{W},\boldsymbol{\Sigma}_0)$:

$$N \times F_{\text{ML}}(\mathbf{W},\boldsymbol{\Sigma}_0) = -2\ln LR$$
$$N \times F_{\text{ML}}(\mathbf{W},\boldsymbol{\Sigma}_0) = -2\ln \frac{L_0}{L_1}$$
$$N \times F_{\text{ML}}(\mathbf{W},\boldsymbol{\Sigma}_0) = -2\ln L_0 + 2\ln L_1$$
$$N \times F_{\text{ML}}(\mathbf{W},\boldsymbol{\Sigma}_0) = N\ln|\boldsymbol{\Sigma}_0| + Np\ln(2\pi) + Ntr\left(\boldsymbol{\Sigma}_0^{-1}\mathbf{W}\right) - N\ln|\mathbf{W}| - Np\ln(2\pi) - Ntr\left(\mathbf{W}^{-1}\mathbf{W}\right)$$

Divide through by N.

$$F_{\text{ML}}(\mathbf{W},\boldsymbol{\Sigma}_0) = \ln|\boldsymbol{\Sigma}_0| + \cancel{p\ln(2\pi)} + tr\left(\boldsymbol{\Sigma}_0^{-1}\mathbf{W}\right) - \ln|\mathbf{W}| - \cancel{p\ln(2\pi)} - tr\left(\mathbf{W}^{-1}\mathbf{W}\right)$$
$$F_{\text{ML}}(\mathbf{W},\boldsymbol{\Sigma}_0) = \ln|\boldsymbol{\Sigma}_0| - \ln|\mathbf{W}| + tr\left(\boldsymbol{\Sigma}_0^{-1}\mathbf{W}\right) - tr\left(\mathbf{W}^{-1}\mathbf{W}\right)$$
$$F_{\text{ML}}(\mathbf{W},\boldsymbol{\Sigma}_0) = \ln|\boldsymbol{\Sigma}_0| - \ln|\mathbf{W}| + tr\left(\boldsymbol{\Sigma}_0^{-1}\mathbf{W} - \mathbf{W}^{-1}\mathbf{W}\right) \text{ combine traces}$$
$$F_{\text{ML}}(\mathbf{W},\boldsymbol{\Sigma}_0) = \ln|\boldsymbol{\Sigma}_0| - \ln|\mathbf{W}| + tr\left(\boldsymbol{\Sigma}_0^{-1}\mathbf{W} - \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_0\right) \text{ replace } \mathbf{W}^{-1}\mathbf{W} \text{ with } \mathbf{I} = \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_0$$
$$F_{\text{ML}}(\mathbf{W},\boldsymbol{\Sigma}_0) = \ln|\boldsymbol{\Sigma}_0| - \ln|\mathbf{W}| + tr\left[\boldsymbol{\Sigma}_0^{-1}\left(\mathbf{W} - \boldsymbol{\Sigma}_0\right)\right] \text{ distributive law}$$
$$F_{\text{ML}}(\mathbf{W},\boldsymbol{\Sigma}_0) = \ln|\boldsymbol{\Sigma}_0| - \ln|\mathbf{W}| + tr\left[\left(\mathbf{W} - \boldsymbol{\Sigma}_0\right)\boldsymbol{\Sigma}_0^{-1}\right] \text{ b/c } tr(\mathbf{AB}) = tr(\mathbf{BA})$$

An equivalent expression that is commonly seen is:

$$F_{\text{ML}}(\mathbf{W},\boldsymbol{\Sigma}_0) = \ln|\boldsymbol{\Sigma}_0| - \ln|\mathbf{W}| + tr\left(\boldsymbol{\Sigma}_0^{-1}\mathbf{W} - \mathbf{W}^{-1}\mathbf{W}\right)$$
$$F_{\text{ML}}(\mathbf{W},\boldsymbol{\Sigma}_0) = \ln|\boldsymbol{\Sigma}_0| - \ln|\mathbf{W}| + tr\left(\boldsymbol{\Sigma}_0^{-1}\mathbf{W} - \mathbf{I}\right)$$
$$F_{\text{ML}}(\mathbf{W},\boldsymbol{\Sigma}_0) = \ln|\boldsymbol{\Sigma}_0| - \ln|\mathbf{W}| + tr\left(\boldsymbol{\Sigma}_0^{-1}\mathbf{W}\right) - tr(\mathbf{I})$$
$$F_{\text{ML}}(\mathbf{W},\boldsymbol{\Sigma}_0) = \ln|\boldsymbol{\Sigma}_0| - \ln|\mathbf{W}| + tr\left(\boldsymbol{\Sigma}_0^{-1}\mathbf{W}\right) - p$$

If we elect to include a mean structure in the model, we cannot pretend that $\boldsymbol{\mu}_0 = \bar{\mathbf{x}}$, so:

$$-2\ln L_0 = Np\ln(2\pi) + N\ln|\boldsymbol{\Sigma}_0| + \sum_{i=1}^{N}(\mathbf{x}_i - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_0)$$
$$-2\ln L_0 = Np\ln(2\pi) + N\ln|\boldsymbol{\Sigma}_0| + NN^{-1}\sum_{i=1}^{N} tr\left\{(\mathbf{x}_i - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_0)\right\}$$
$$-2\ln L_0 = Np\ln(2\pi) + N\ln|\boldsymbol{\Sigma}_0| + Ntr\left\{\boldsymbol{\Sigma}_0^{-1}N^{-1}\sum_{i=1}^{N}(\mathbf{x}_i - \boldsymbol{\mu}_0)(\mathbf{x}_i - \boldsymbol{\mu}_0)'\right\}$$
$$-2\ln L_0 = Np\ln(2\pi) + N\ln|\boldsymbol{\Sigma}_0| + Ntr\left\{\boldsymbol{\Sigma}_0^{-1}N^{-1}\sum_{i=1}^{N}(\mathbf{x}_i - \bar{\mathbf{x}} + \bar{\mathbf{x}} - \boldsymbol{\mu}_0)(\mathbf{x}_i - \bar{\mathbf{x}} + \bar{\mathbf{x}} - \boldsymbol{\mu}_0)'\right\}$$

$$-2\ln L_0 = Np\ln(2\pi) + N\ln|\mathbf{\Sigma}_0| + Ntr\left\{\mathbf{\Sigma}_0^{-1}N^{-1}\sum_{i=1}^{N}\left((\mathbf{x}_i-\bar{\mathbf{x}})+(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)\right)\left((\mathbf{x}_i-\bar{\mathbf{x}})'+(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)'\right)\right\}$$

$$-2\ln L_0 = Np\ln(2\pi) + N\ln|\mathbf{\Sigma}_0| + Ntr\left\{\mathbf{\Sigma}_0^{-1}N^{-1}\sum_{i=1}^{N}(\mathbf{x}_i-\bar{\mathbf{x}})(\mathbf{x}_i-\bar{\mathbf{x}})'\right\} + Ntr\left\{\mathbf{\Sigma}_0^{-1}N^{-1}\sum_{i=1}^{N}(\mathbf{x}_i-\bar{\mathbf{x}})(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)'\right\}$$

$$+Ntr\left\{\mathbf{\Sigma}_0^{-1}N^{-1}\sum_{i=1}^{N}(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)(\mathbf{x}_i-\bar{\mathbf{x}})'\right\} + Ntr\left\{\mathbf{\Sigma}_0^{-1}N^{-1}\sum_{i=1}^{N}(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)'\right\}$$

$$-2\ln L_0 = Np\ln(2\pi) + N\ln|\mathbf{\Sigma}_0| + Ntr\left\{\mathbf{\Sigma}_0^{-1}\mathbf{W}\right\} + Ntr\left\{\mathbf{\Sigma}_0^{-1}N^{-1}\cancel{\sum_{i=1}^{N}(\mathbf{x}_i-\bar{\mathbf{x}})(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)'}\right\}$$

$$+Ntr\left\{\mathbf{\Sigma}_0^{-1}N^{-1}\cancel{\sum_{i=1}^{N}(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)(\mathbf{x}_i-\bar{\mathbf{x}})'}\right\} + Ntr\left\{\mathbf{\Sigma}_0^{-1}N^{-1}\sum_{i=1}^{N}(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)'\right\}$$

$$-2\ln L_0 = Np\ln(2\pi) + N\ln|\mathbf{\Sigma}_0| + Ntr\left\{\mathbf{\Sigma}_0^{-1}\mathbf{W}\right\} + Ntr\left\{\mathbf{\Sigma}_0^{-1}N^{-1}N(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)'\right\}$$

$$-2\ln L_0 = Np\ln(2\pi) + N\ln|\mathbf{\Sigma}_0| + Ntr\left\{\mathbf{\Sigma}_0^{-1}\mathbf{W}\right\} + Ntr\left\{(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)'\mathbf{\Sigma}_0^{-1}(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)\right\}$$

$$-2\ln L_0 = Np\ln(2\pi) + N\ln|\mathbf{\Sigma}_0| + Ntr\left\{\mathbf{\Sigma}_0^{-1}\mathbf{W}\right\} + N(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)'\mathbf{\Sigma}_0^{-1}(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)$$

As for the saturated model,

$$-2\ln L_1 = Np\ln(2\pi) + N\ln|\mathbf{W}| + \sum_{i=1}^{N}(\mathbf{x}_i-\bar{\mathbf{x}})'\mathbf{W}^{-1}(\mathbf{x}_i-\bar{\mathbf{x}})$$

$$-2\ln L_1 = Np\ln(2\pi) + N\ln|\mathbf{W}| + NN^{-1}\sum_{i=1}^{N}(\mathbf{x}_i-\bar{\mathbf{x}})'\mathbf{W}^{-1}(\mathbf{x}_i-\bar{\mathbf{x}})$$

$$-2\ln L_1 = Np\ln(2\pi) + N\ln|\mathbf{W}| + Ntr\left(\mathbf{W}^{-1}\mathbf{W}\right)$$

$$-2\ln L_1 = Np\ln(2\pi) + N\ln|\mathbf{W}| + Np$$

From these it can be seen that the discrepancy function *including* mean structures will have an additional term from the proposed model's likelihood:

$$F_{ML}(\mathbf{W},\mathbf{\Sigma}_0) = \ln|\mathbf{\Sigma}_0| - \ln|\mathbf{W}| + tr\left\{\mathbf{\Sigma}_0^{-1}\mathbf{W}\right\} - p + (\bar{\mathbf{x}}-\boldsymbol{\mu}_0)'\mathbf{\Sigma}_0^{-1}(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)$$

Conveniently, $N \times F_{ML} \sim \chi^2_{df}$.

Discrepancy based on the Wishart distribution

One expression of the Wishart likelihood (instead involving the unbiased estimate of $\boldsymbol{\Sigma}$, the sample covariance matrix $\mathbf{S}$) is:[2]

$$C\left|\mathbf{S}\right|^{\frac{1}{2}(n-p-1)}\left|\boldsymbol{\Sigma}_0\right|^{-\frac{n}{2}}\exp\left(-\frac{n}{2}tr\left\{\boldsymbol{\Sigma}_M^{-1}\mathbf{S}\right\}\right)$$

where $n = N-1$ and C is a constant that does not depend on the model parameters. Algebra similar to that above will yield a discrepancy function identical to the one derived using the MVN likelihood. First, the likelihood for the proposed model is:

$$-2\ln L_0 = -2\ln\left(C\left|\mathbf{S}\right|^{\frac{1}{2}(n-p-1)}\left|\boldsymbol{\Sigma}_0\right|^{-\frac{n}{2}}\exp\left(-\frac{n}{2}tr\left\{\boldsymbol{\Sigma}_0^{-1}\mathbf{S}\right\}\right)\right)$$
$$-2\ln L_0 = -2\ln(C) - (n-p-1)\ln\left|\mathbf{S}\right| + n\ln\left|\boldsymbol{\Sigma}_0\right| + ntr\left\{\boldsymbol{\Sigma}_0^{-1}\mathbf{S}\right\}$$
$$-2\ln L_0 = -2\ln(C) - (n-p-1)\ln\left|\mathbf{S}\right| + n\ln\left|\boldsymbol{\Sigma}_0\right| + ntr\left\{\boldsymbol{\Sigma}_0^{-1}\mathbf{S}\right\}$$

The likelihood for the saturated model is:

$$-2\ln L_1 = -2\ln\left(C\left|\mathbf{S}\right|^{\frac{1}{2}(n-p-1)}\left|\mathbf{S}\right|^{-\frac{n}{2}}\exp\left(-\frac{n}{2}tr\left\{\mathbf{S}^{-1}\mathbf{S}\right\}\right)\right)$$
$$-2\ln L_1 = -2\ln(C) - (n-p-1)\ln\left|\mathbf{S}\right| + n\ln\left|\mathbf{S}\right| + np$$
$$-2\ln L_1 = -2\ln(C) + (p+1)\ln\left|\mathbf{S}\right| + np$$

Finally,

$$(N-1)\times F_{\mathrm{ML}}(\mathbf{S},\boldsymbol{\Sigma}_0) = -2\ln LR$$
$$(N-1)\times F_{\mathrm{ML}}(\mathbf{S},\boldsymbol{\Sigma}_0) = -2\ln L_0 + 2\ln L_1$$
$$(N-1)\times F_{\mathrm{ML}}(\mathbf{S},\boldsymbol{\Sigma}_0) = \cancel{-2\ln(C)} - (n-p-1)\ln\left|\mathbf{S}\right| + n\ln\left|\boldsymbol{\Sigma}_0\right| + ntr\left\{\boldsymbol{\Sigma}_0^{-1}\mathbf{S}\right\} + \cancel{2\ln(C)} - (p+1)\ln\left|\mathbf{S}\right| - np$$
$$(N-1)\times F_{\mathrm{ML}}(\mathbf{S},\boldsymbol{\Sigma}_0) = n\ln\left|\boldsymbol{\Sigma}_0\right| - n\ln\left|\mathbf{S}\right| + ntr\left\{\boldsymbol{\Sigma}_0^{-1}\mathbf{S}\right\} - np$$
$$F_{\mathrm{ML}}(\mathbf{S},\boldsymbol{\Sigma}_0) = \ln\left|\boldsymbol{\Sigma}_0\right| - \ln\left|\mathbf{S}\right| + tr\left\{\boldsymbol{\Sigma}_0^{-1}\mathbf{S}\right\} - p$$

Note that if the Wishart likelihood is used as a basis for deriving F_{ML}, then $(N-1)\times F_{\mathrm{ML}} \sim \chi^2_{df}$.

---

[2] This expression differs slightly from one given in many multivariate texts, $C^*\left|\mathbf{A}\right|^{\frac{1}{2}(n-p-1)}\left|\boldsymbol{\Sigma}_0\right|^{-n/2}\exp(-\frac{1}{2}tr\{\boldsymbol{\Sigma}_0^{-1}\mathbf{A}\})$, where $\mathbf{A}$ is a cross-product matrix equal to $n\mathbf{S}$ and $C^* = Cn^{-\frac{1}{2}(n-p-1)}$.

References

Anderson, T. W., & Amemiya, Y. (1988). The asymptotic normal distribution of estimators in factor analysis under general conditions. *The Annals of Statistics*, *16*, 759-771.

[3]Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons, Inc.

Browne, M. W., & Arminger, G. (1994). Specification and estimation of mean- and covariance-structure models. In Arminger, G., Clogg, C. C., & Sobel, M. E. (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (pp. 185-249). New York: Plenum Press.

Browne, M. W., Cudeck, R., Tateneni, K., & Mels, G. (2008). *CEFA: Comprehensive exploratory factor analysis*. Available: http://faculty.psy.ohio-state.edu/browne/

Muthén, B. O. (2004). *Mplus technical appendices*. Los Angeles, CA: Muthén & Muthén. Available: http://www.statmodel.com/

[3] See Appendices 4A and 4B, pp. 131-135.