Test for Covariances

Max Turgeon

STAT 7200–Multivariate Statistics

Objectives

- Review general theory of likelihood ratio tests
- Tests for structured covariance matrices
- Test for equality of multiple covariance matrices

Likelihood ratio tests i

- We will build our tests for covariances using likelihood ratios.
  - Therefore, we quickly review the asymptotic theory for regular models.
- Let $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ be a random sample from a density p_θ with parameter $\theta \in \mathbb{R}^d$.
- We are interested in the following hypotheses:

$$H_0 : \theta \in \Theta_0, \qquad H_1 : \theta \in \Theta_1,$$

where $\Theta_i \subseteq \mathbb{R}^d$.

Likelihood ratio tests ii

- Let $L(\theta) = \prod_{i=1}^n p_\theta(\mathbf{Y}_i)$ be the likelihood, and define the likelihood ratio

$$\Lambda = \frac{\max_{\theta\in\Theta_0} L(\theta)}{\max_{\theta\in\Theta_0\cup\Theta_1} L(\theta)}.$$

- **Recall**: we reject the null hypothesis H_0 for small values of Λ.

Likelihood ratio tests iii

Theorem (Van der Wandt, Chapter 16)
Assume Θ_0, Θ_1 are *locally linear*. Under regularity conditions on p_θ, we have

$$-2 \log \Lambda \to \chi^2(k),$$

where k is the difference in the number of free parameters between the null model Θ_0 and the unrestricted model $\Theta_0 \cup \Theta_1$.

- Therefore, in practice, we need to count the number of free parameters in each model and hope the sample size n is large enough.

- We are going to look at several tests for structured covariance matrix.
- Throughout, we assume $\mathbf{Y}_1, \ldots, \mathbf{Y}_n \sim N_p(\mu, \Sigma)$ with Σ positive definite.
    - Like other exponential families, the multivariate normal distribution satisfies the regularity conditions of the theorem above.
    - Being positive definite implies that the unrestricted parameter space is *locally linear*, i.e. we are staying away from the boundary where Σ is singular.

Tests for structured covariance matrices ii

- A few important observations about the unrestricted model:
  - The number of free parameters is equal to the number of entries on and above the diagonal of Σ, which is $p(p+1)/2$.
  - The sample mean $\bar{\mathbf{Y}}$ maximises the likelihood **independently of the structure of Σ**.
  - The maximised likelihood for the unrestricted model is given by

$$L(\hat{\mathbf{Y}}, \hat{\Sigma}) = \frac{\exp(-np/2)}{(2\pi)^{np/2}|\hat{\Sigma}|^{n/2}}.$$

Specified covariance structure i

- We will start with the simplest hypothesis test:

$$H_0 : \Sigma = \Sigma_0.$$

- Note that there is no free parameter in the null model.
- Write $V = n\hat{\Sigma}$. Recall that we have

$$L(\hat{\mathbf{Y}}, \Sigma) = (2\pi)^{-np/2}|\Sigma|^{-n/2}\exp\left(-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}V)\right).$$

Specified covariance structure ii

- Therefore, the likelihood ratio is given by

$$\begin{aligned}\Lambda &= \frac{(2\pi)^{-np/2}|\Sigma_0|^{-n/2}\exp\left(-\frac{1}{2}\mathrm{tr}(\Sigma_0^{-1}V)\right)}{\exp(-np/2)(2\pi)^{-np/2}|\hat{\Sigma}|^{-n/2}} \\ &= \frac{|\Sigma_0|^{-n/2}\exp\left(-\frac{1}{2}\mathrm{tr}(\Sigma_0^{-1}V)\right)}{\exp(-np/2)|n^{-1}V|^{-n/2}} \\ &= \left(\frac{e}{n}\right)^{np/2}|\Sigma_0^{-1}V|^{n/2}\exp\left(-\frac{1}{2}\mathrm{tr}(\Sigma_0^{-1}V)\right).\end{aligned}$$

- In particular, if $\Sigma_0 = I_p$, we get

$$\Lambda = \left(\frac{e}{n}\right)^{np/2}|V|^{n/2}\exp\left(-\frac{1}{2}\mathrm{tr}(V)\right).$$

Example i

```r
library(tidyverse)
# Winnipeg avg temperature
url <- paste0("https://maxturgeon.ca/w20-stat7200/",
              "winnipeg_temp.csv")
dataset <- read.csv(url)
dataset[1:3,1:3]
```

```
##   temp_2010 temp_2011  temp_2012
## 1 -25.57500 -16.25417  -6.379167
## 2 -26.06250 -18.39583 -12.925000
## 3 -20.56667 -19.45833  -5.791667
```

Example ii

```r
n <- nrow(dataset)
p <- ncol(dataset)

V <- (n - 1)*cov(dataset)
```

```r
# Diag = 14^2
# Corr = 0.8
Sigma0 <- diag(0.8, nrow = p)
diag(Sigma0) <- 1
Sigma0 <- 14^2*Sigma0
Sigma0_invXV <- solve(Sigma0, V)
```

Example iii

```
lrt <- 0.5*n*p*(1 - log(n))
lrt <- lrt + 0.5*n*log(det(Sigma0_invXV))
lrt <- lrt - 0.5*sum(diag(Sigma0_invXV))
lrt <- -2*lrt
```

```
df <- choose(p + 1, 2)
c(lrt, qchisq(0.95, df))
```

```
## [1] 5631.63409   73.31149
```

Test for sphericity i

- *Sphericity* means the different components of $\mathbf{Y}$ are **uncorrelated** and have the **same variance**.
    - In other words, we are looking at the following null hypothesis:

$$H_0 : \Sigma = \sigma^2 I_p, \qquad \sigma^2 > 0.$$

    - Note that there is one free parameter.
- We have

$$\begin{aligned} L(\hat{\mathbf{Y}}, \sigma^2 I_p) &= (2\pi)^{-np/2} |\sigma^2 I_p|^{-n/2} \exp\left(-\frac{1}{2} \mathrm{tr}((\sigma^2 I_p)^{-1} V) \right) \\ &= (2\pi\sigma^2)^{-np/2} \exp\left(-\frac{1}{2\sigma^2} \mathrm{tr}(V) \right). \end{aligned}$$

Test for sphericity ii

- Taking the derivative of the logarithm and setting it equal to zero, we find that $L(\hat{\mathbf{Y}}, \sigma^2 I_p)$ is maximised when

$$\widehat{\sigma^2} = \frac{\mathrm{tr} V}{np}.$$

- We then get

$$\begin{aligned} L(\hat{\mathbf{Y}}, \widehat{\sigma^2} I_p) &= (2\pi\widehat{\sigma^2})^{-np/2} \exp\left(-\frac{1}{2\widehat{\sigma^2}}\mathrm{tr}(V)\right) \\ &= (2\pi)^{-np/2}\left(\frac{\mathrm{tr} V}{np}\right)^{-np/2} \exp\left(-\frac{np}{2}\right). \end{aligned}$$

Test for sphericity iii

- Therefore, we have

$$\begin{aligned}
\Lambda &= \frac{(2\pi)^{-np/2}\left(\frac{\mathrm{tr}V}{np}\right)^{-np/2}\exp\left(-\frac{np}{2}\right)}{\exp(-np/2)(2\pi)^{-np/2}|\hat{\Sigma}|^{-n/2}} \\
&= \frac{\left(\frac{\mathrm{tr}V}{np}\right)^{-np/2}}{|n^{-1}V|^{-n/2}} \\
&= \left(\frac{|V|}{(\mathrm{tr}V/p)^p}\right)^{n/2}.
\end{aligned}$$

Example (cont'd) i

```
lrt <- -2*0.5*n*(log(det(V)) - p*log(mean(diag(V))))
df <- choose(p + 1, 2) - 1

c(lrt, qchisq(0.95, df))
```

```
## [1] 5630.79458   72.15322
```

- Recall that we have

$$\Lambda = \left(\frac{|V|}{(\mathrm{tr}V/p)^p}\right)^{n/2}.$$

- We can rewrite this as follows: let $l_1 \geq \cdots \geq l_p$ be the eigenvalues of V. We have

$$\begin{aligned}
\Lambda^{2/n} &= \frac{|V|}{(\mathrm{tr}V/p)^p} \\
&= \frac{\prod_{j=1}^p l_j}{(\frac{1}{p}\sum_{j=1}^p l_j)^p} \\
&= \left(\frac{\prod_{j=1}^p l_j^{1/p}}{\frac{1}{p}\sum_{j=1}^p l_j}\right)^p.
\end{aligned}$$

Test for sphericity (cont'd) ii

- In other words, the modified LRT $\tilde{\Lambda} = \Lambda^{2/n}$ is the ratio of the geometric to the arithmetic mean of the eigenvalues of V (all raised to the power p).
- A result of Srivastava and Khatri gives the *exact* distribution of $\tilde{\Lambda}$:

$$\tilde{\Lambda} = \prod_{j=1}^{p-1} \mathcal{B}\left(\frac{1}{2}(n-j-1), j\left(\frac{1}{2}+\frac{1}{p}\right)\right).$$

Example (cont'd) i

```r
B <- 1000
df1 <- 0.5*(n - seq_len(p-1) - 1)
df2 <- seq_len(p-1)*(0.5 + 1/p)

# Critical values
dist <- replicate(B, {
  prod(rbeta(p-1, df1, df2))
  })
```

Example (cont'd) ii

```
# Test statistic
decomp <- eigen(V, symmetric = TRUE, only.values = TRUE)
ar_mean <- mean(decomp$values)
geo_mean <- exp(mean(log(decomp$values)))

lrt_mod <- (geo_mean/ar_mean)^p

c(lrt_mod, quantile(dist, 0.95))
```

```
##                       95%
## 1.181561e-07 8.967361e-01
```

Test for independence i

- Decompose $\mathbf{Y}_i$ into k blocks:

$$\mathbf{Y}_i = (\mathbf{Y}_{1i}, \ldots, \mathbf{Y}_{ki}),$$

  where $\mathbf{Y}_{ji} \sim N_{p_j}(\mu_j, \Sigma_{jj})$ and $\sum_{j=1}^k p_j = p$.
- This induces a decomposition on Σ and V:

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \cdots & \Sigma_{1k} \\ \vdots & \ddots & \vdots \\ \Sigma_{k1} & \cdots & \Sigma_{kk} \end{pmatrix}, \qquad V = \begin{pmatrix} V_{11} & \cdots & V_{1k} \\ \vdots & \ddots & \vdots \\ V_{k1} & \cdots & V_{kk} \end{pmatrix}.$$

Test for independence ii

- We are interested in testing for independence between the different blocks $\mathbf{Y}_{1i}, \ldots, \mathbf{Y}_{ki}$. This equivalent to

$$H_0 : \Sigma = \begin{pmatrix} \Sigma_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Sigma_{kk} \end{pmatrix}.$$

    - Note that there are $\sum_{j=1}^{k} p_j(p_j + 1)/2$ free parameters.
- Under the null hypothesis, the likelihood can be decomposed into k likelihoods that can be maximised independently.

- This gives us

$$\max L(\hat{\mathbf{Y}}, \Sigma) = \prod_{j=1}^{k} \frac{\exp(-np_j/2)}{(2\pi)^{np_j/2}|\widehat{\Sigma_{jj}}|^{n/2}}$$
$$= \frac{\exp(-np/2)}{(2\pi)^{np/2}\prod_{j=1}^{k}|\widehat{\Sigma_{jj}}|^{n/2}}.$$

- Putting this together, we conclude that

$$\Lambda = \left(\frac{|V|}{\prod_{j=1}^{k}|V_{jj}|}\right)^{n/2}.$$

Example i

```r
url <- paste0("https://maxturgeon.ca/w20-stat7200/",
              "blue_data.csv")
blue_data <- read.csv(url)
names(blue_data)
```

```
## [1] "NumSold"     "Price"       "AdvCost"     "SalesAs
```

```r
dim(blue_data)
```

```
## [1] 10  4
```

Example ii

```
# Let's test for independence between
# all four variables
n <- nrow(blue_data)
p <- ncol(blue_data)

V <- (n-1)*cov(blue_data)
lrt <- -2*(log(det(V)) - sum(log(diag(V))))
```

```
df <- choose(p + 1, 2) - p
c(lrt, qchisq(0.95, df))
```

Example iii

```
## [1]  5.635124 12.591587
```

```r
lrt > qchisq(0.95, df)
```

```
## [1] FALSE
```

Test for equality of covariances i

- We now look at a different setting: assume that we collected K independent random samples from (potentially) different p-dimensional multivariate normal distributions:

$$\mathbf{Y}_{1k}, \ldots, \mathbf{Y}_{n_k k} \sim N_p(\mu_k, \Sigma_k), \quad k = 1, \ldots, K.$$

- We are interested in the null hypothesis that all Σ_k are equal to some unknown Σ:

$$H_0 : \Sigma_k = \Sigma, \quad \text{for all } k = 1, \ldots, K.$$

Test for equality of covariances ii

- First, note that since the samples are independent, the full likelihood is the product of the likelihoods for each sample:

$$L(\mu_1, \ldots, \mu_K, \Sigma_1, \ldots, \Sigma_K) = \prod_{k=1}^{K} L(\mu_k, \Sigma_k).$$

- Therefore, over the unrestricted model, the maximum likelihood estimators are

$$(\bar{\mathbf{Y}}_k, \hat{\Sigma}_k).$$

  - Note that the number of free parameters over the unrestricted model is $kp(p+1)/2$.

Test for equality of covariances iii

- Now, over the null model, the full likelihood is still maximised when $\mu_k = \bar{\mathbf{Y}}_k$. Hence, we get

$$\begin{aligned}
L(\bar{\mathbf{Y}}_1, \ldots, \bar{\mathbf{Y}}_K, \Sigma, \ldots, \Sigma) &= \prod_{k=1}^K L(\bar{\mathbf{Y}}_K, \Sigma) \\
&= \prod_{k=1}^K (2\pi)^{-n_k p/2} |\Sigma|^{-n_k/2} \exp\left(-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}V_k)\right),
\end{aligned}$$

  where $V_k = n_k\hat{\Sigma}_k$.
- Writing $n = \sum_{k=1}^K n_k$ and $V = \sum_{k=1}^K V_k$, we get

$$\begin{aligned}
&L(\bar{\mathbf{Y}}_1, \ldots, \bar{\mathbf{Y}}_K, \Sigma, \ldots, \Sigma) = \\
&\quad = (2\pi)^{-np/2} |\Sigma|^{-n/2} \exp\left(-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}V)\right).
\end{aligned}$$

Test for equality of covariances iv

- This is the same expression as the one we would get by pooling all the samples together. Therefore, the maximum likelihood estimate is
$$\hat{\Sigma} = \frac{1}{n} V.$$
    - Note that under the null model, there are $p(p+1)/2$ free parameters.

- We can now compute the likelihood ratio:

$$\begin{aligned}
\Lambda &= \frac{L(\bar{\mathbf{Y}}_1, \ldots, \bar{\mathbf{Y}}_K, \hat{\Sigma}, \ldots, \hat{\Sigma})}{L(\bar{\mathbf{Y}}_1, \ldots, \bar{\mathbf{Y}}_K, \hat{\Sigma}_1, \ldots, \hat{\Sigma}_K)} \\
&= \frac{(2\pi)^{-np/2} \exp(-np/2) |\hat{\Sigma}|^{-n/2}}{\prod_{k=1}^K (2\pi)^{-n_k p/2} \exp(-n_k p/2) |\hat{\Sigma}_k|^{-n_k/2}} \\
&= \frac{(2\pi)^{-np/2} \exp(-np/2) |\hat{\Sigma}|^{-n/2}}{(2\pi)^{-np/2} \exp(-np/2) \prod_{k=1}^K |\hat{\Sigma}_k|^{-n_k/2}} \\
&= \frac{|\hat{\Sigma}|^{-n/2}}{\prod_{k=1}^K |\hat{\Sigma}_k|^{-n_k/2}}.
\end{aligned}$$

- In other words, the likelihood ratio test compares the generalized variance of the *pooled covariance* with the product of the generalized variances of the *individuals covariances*.
- From the general theory of LRTs, we get

$$-2\log\Lambda \approx \chi^2\left(\frac{(K-1)p(p+1)}{2}\right).$$

Test for equality of covariances vii

```r
## Example on producing plastic film
## from Krzanowski (1998, p. 381)
tear <- c(6.5, 6.2, 5.8, 6.5, 6.5, 6.9, 7.2,
          6.9, 6.1, 6.3, 6.7, 6.6, 7.2, 7.1,
          6.8, 7.1, 7.0, 7.2, 7.5, 7.6)
gloss <- c(9.5, 9.9, 9.6, 9.6, 9.2, 9.1, 10.0,
           9.9, 9.5, 9.4, 9.1, 9.3, 8.3, 8.4,
           8.5, 9.2, 8.8, 9.7, 10.1, 9.2)
opacity <- c(4.4, 6.4, 3.0, 4.1, 0.8, 5.7, 2.0,
             3.9, 1.9, 5.7, 2.8, 4.1, 3.8, 1.6,
             3.4, 8.4, 5.2, 6.9, 2.7, 1.9)
```

Test for equality of covariances viii

```
Y <- cbind(tear, gloss, opacity)
Y_low <- Y[1:10,]
Y_high <- Y[11:20,]
n <- nrow(Y); p <- ncol(Y); K <- 2
n1 <- n2 <- nrow(Y_low)
```

Test for equality of covariances ix

```r
Sig_low <- (n1 - 1)*cov(Y_low)/n1
Sig_high <- (n2 - 1)*cov(Y_high)/n2
Sig_pool <- (n1*Sig_low + n2*Sig_high)/n

c("pool" = log(det(Sig_pool)),
  "low" = log(det(Sig_low)),
  "high" = log(det(Sig_high)))
```

```
##      pool       low      high
## -2.524791 -3.265178 -2.329143
```

Test for equality of covariances x

```r
lrt <- n*log(det(Sig_pool)) -
  n1*log(det(Sig_low)) -
  n2*log(det(Sig_high))
df <- (K - 1)*choose(p + 1, 2)
c(lrt, qchisq(0.95, df))
```

```
## [1]  5.447396 12.591587
```

- There are a few ways to get a better approximation of the null distribution of Λ. First, note that we can rewrite it as

$$\Lambda = \frac{\prod_{k=1}^K |V_k|^{n_k/2}}{|V|^{n/2}} \frac{n^{pn/2}}{\prod_{k=1}^K n_k^{pn_k/2}}.$$

- We can create an *unbiased* test (i.e. it has the correct asymptotic expectation) by replacing n_k by $n_k - 1$ and n with $n - K$:

$$\Lambda^* = \frac{\prod_{k=1}^K |V_k|^{(n_k-1)/2}}{|V|^{(n-K)/2}} \frac{(n-K)^{p(n-K)/2}}{\prod_{k=1}^K (n_k-1)^{p(n_k-1)/2}}.$$

  - This is equivalent to replacing $\hat{\Sigma}_k$ by the sample covariances S_k.

- Note that we still have the same asymptotic result:

$$-2\log\Lambda^* \approx \chi^2\left(\frac{(K-1)p(p+1)}{2}\right).$$

- Box showed that you can further improve the approximation by multiplying the test statistic by a constant. Set

$$u = \left(\sum_{k=1}^K \frac{1}{n_k - 1} - \frac{1}{n-K}\right)\left(\frac{2p^2+3p-1}{6(p+1)(K-1)}\right).$$

- Then we have

$$-2(1-u)\log\Lambda^* \approx \chi^2\left(\frac{(K-1)p(p+1)}{2}\right).$$

Example (cont'd) i

```r
S_low <- cov(Y_low)
S_high <- cov(Y_high)
S_pool <- ((n1 - 1)*S_low + (n2 - 1)*S_high)/(n - K)

lrt2 <- (n - K)*log(det(S_pool)) -
  (n1 - 1)*log(det(S_low)) -
  (n2 - 1)*log(det(S_high))

c(lrt, lrt2, qchisq(0.95, df))
```

```
## [1]  5.447396  4.902657 12.591587
```

Example (cont'd) ii

```
u <- (2*p^2 + 3*p - 1)/(6*(p + 1)*(K - 1))
u <- u * ((n1 - 1)^{-1} + (n2 - 1)^{-1} - (n - K)^{-1})
lrt3 <- lrt2*(1 - u)

c(lrt, lrt2, lrt3, qchisq(0.95, df))
```

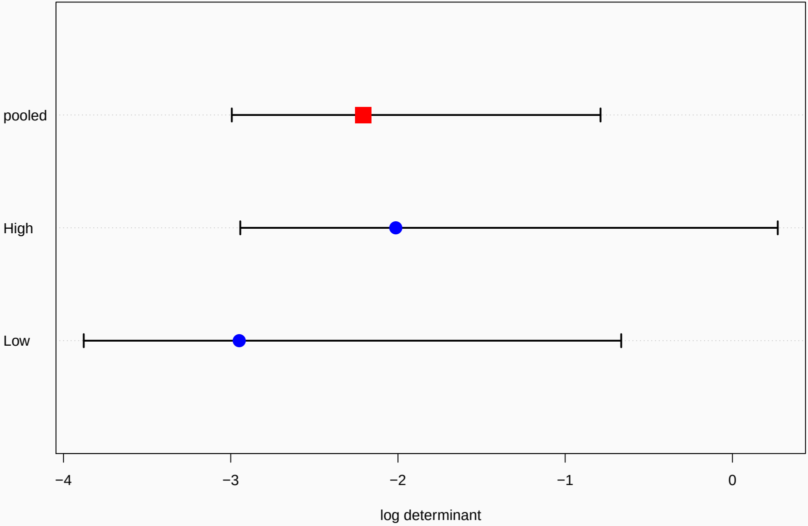
```
## [1]  5.447396  4.902657  4.017455 12.591587
```

Visualization i

```
# You can also visualize the covariances----
library(heplots)
rate <- gl(K, 10, labels = c("Low", "High"))
boxm_res <- boxM(Y, rate)

# You can plot the log generalized variances
# The plot function adds 95% CI
plot(boxm_res)
```

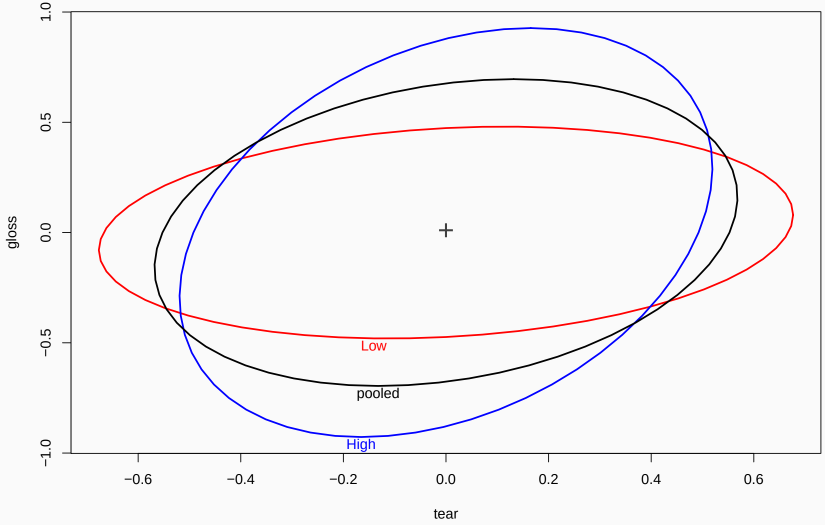
Visualization ii



Visualization iii

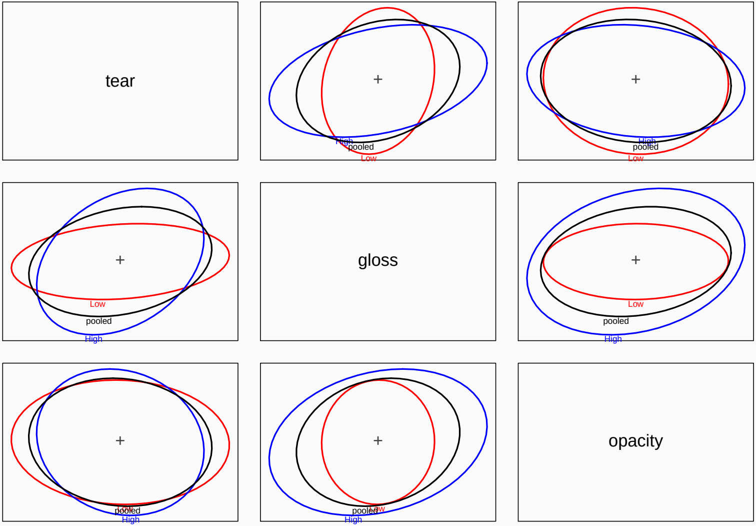
```
# Finally you can also plot the ellipses
# as a way to compare the covariances
covEllipses(Y, rate, center = TRUE,
            label.pos = 'bottom')
```

Visualization iv



```r
# Or all pairwise comparisons together
covEllipses(Y, rate, center = TRUE,
            label.pos = 'bottom',
            variables = 1:3)
```

Visualization vi



Asymptotic expansions for likelihood ratio tests i

- Box's correction of the LRT for equality of covariances is part of a general theory of *asymptotic expansions* for LRTs.
    - The frameword allows for approximations of the null distribution of some LRTs to any degrees of accuracy.
- We won't go into the details of such expansions, but we will look at one example.
    - If you want more details, see this: `https://maxturgeon.ca/w20-stat7200/test-sphericity-details.pdf`

Asymptotic expansions for likelihood ratio tests ii

- In the context of the test for sphericity, the approximation result looks like this:

$$-2\left(\frac{6p(n-1)-(2p^2+p+2)}{6pn}\right)\log\Lambda \approx \chi^2\left(\frac{1}{2}p(p+1)-1\right),$$

  where Λ is the likelihood ratio.
- This is known also known as *Bartlett's correction*.
    - Note that we are correcting both the test statistic (by multiplying by a positive constant) and the degrees of freedom (we lose one degree of freedom).

Simulation i

```r
set.seed(7200)

# Simulation parameters
n <- 10
p <- 2
B <- 1000
```

Simulation ii

```r
# Generate data
lrt_dist <- replicate(B, {
  Y <- matrix(rnorm(n*p), ncol = p)
  V <- crossprod(Y)
  # log Lambda
  0.5*n*(log(det(V)) - p*log(mean(diag(V))))
})
```

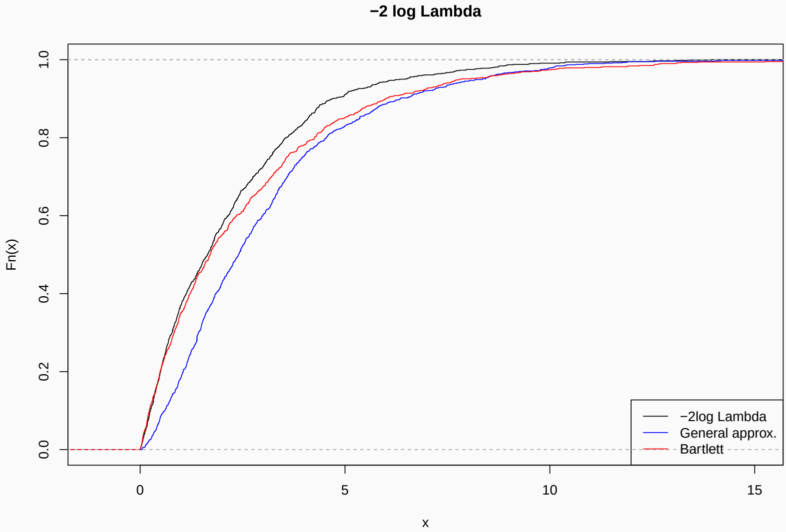
```r
# General asymptotic result
df <- choose(p + 1, 2)
general_chisq <- rchisq(B, df = df)
```

```r
# Bartlett's correction
df <- choose(p + 1, 2) - 1
const <- (6*p*(n-1) - (2*p^2 + p + 2))/(6*p*n)
bartlett_chisq <- rchisq(B, df = df)/const
```

Simulation iv

```
# Plot empirical CDFs
plot(ecdf(-2*lrt_dist), main = "-2 log Lambda")
lines(ecdf(general_chisq), col = 'blue')
lines(ecdf(bartlett_chisq), col = 'red')
legend('bottomright',
       legend = c("-2log Lambda", "General approx.",
                  "Bartlett"),
       lty = 1, col = c('black', 'blue', 'red'))
```

Simulation v



Sketch of a proof i

- Here is an outline of how you could get such an approximation:
- First, we can compute the moments of the likelihood ratio: given h, we have

$$E\left(\Lambda^{2h/n}\right) = p^{ph}\frac{\Gamma\left(\frac{1}{2}(n-1)p\right)}{\Gamma\left(\frac{1}{2}(n-1)p+ph\right)}\frac{\Gamma_p\left(\frac{1}{2}(n-1)+h\right)}{\Gamma_p\left(\frac{1}{2}(n-1)\right)}.$$

- Next, we can use this expression to get an expression for the characteristic function of $\rho M = -2\rho\log\Lambda^{(n-1)/n}$:

$$\varphi_{\rho M}(t) = E(\exp(it\rho M)) = E\left(\Lambda^{-2it\rho(n-1)/n}\right).$$

Sketch of a proof ii

- Therefore, if we take $h = -it\rho(n-1)$, we can see that the characteristic function $\varphi_{\rho M}(t)$ is a product of gamma functions.
- The *cumulant function*, which is the logarithm of the characteristic function, is therefore a sum of logarithms of gamma functions.
- **Why do we care?** We can use Stirling's approximation to approximate the logarithm of gamma functions to any degree of precision.
- This approximation of the cumulant function gives rise to an approximation of the characteristic function. For order 2, we get:

$$\varphi_{\rho M}(t) \approx (1-2it)^{-f/2} + \omega_1 \left((1-2it)^{-(f+2)/2} - (1-2it)^{-f/2}\right).$$

Sketch of a proof iii

- Recall that the characteristic function of $\chi^2(d)$ is $(1 - 2it)^{-d/2}$. Therefore, we can "invert" our approximation of $\varphi_{\rho M}(t)$ to get an approximation of the density and the distribution of ρM.
- Moreover, we can choose ρ in such a way that $\omega_1 = 0$, which gives a chi-square approximation that is **more** accurate than the general asymptotic theory.

Summary

- We built tests for structured covariance matrices using likelihood ratio tests.
- We also built a test for equality of covariance, when we have multiple samples.
- We briefly discussed asymptotic expansions and how they can give rise to better approximations of the likelihood ratio test statistics.