PROOF OF THE COCHRAN's THEOREM

Minatellah Sourour Nehal

February 2026

NOTE:

This proof is the student's personal work. It is inspired by the classical proof of Cochran's theorem presented on **Adam Kashlak**'s YouTube channel. The aim of this proof is to clarify the relationship between linear algebra and statistics underlying this theorem.

The classical version of the Cochran's theorem:

Theorem 1 (Cochran)**.** *let* $Z = (Z_1 \dots Z_n)^T \sim N_n(0, \sigma^2 I_n)$

- *and let:* $k = 1 \dots s$, A_k *symetric idemponent matrices* $n \times n$ *with an* ij *th entry* a_{ij}^k
- *and we have the quadratic form* Q_k *telle que :* $Q_z = Z^T A_k Z = \sum_{ij=1}^n a_{ij}^k Z_i Z_j$
- $\sum_{i=1}^n Z_i^2 = \sum_{k=1}^s Q_k$
- $rank(A_k) = r_k$

So: $\sum_{k=1}^s r_k = n$ *which is the dimension of* A_k ***If and only if****:* Q_k *are independent and* $Q_k/\sigma^2 \sim \chi^2_{r_k}$ *which is the dimension of* A

Explanation:

Before we start the proof we need to clarify each elemnt of theorm and what does it mean statistaclly:

- $Z = (Z_1, \dots, Z_n)^T$: a vector of independent Gaussian random variables.
- A_k: a symmetric matrix defining an orthogonal projection onto a subspace, this means $A_k A_l = 0$ $(k \neq l)$
- $Q_k = Z^T A_k Z$: we introduce them because they are the mathematical objects that measure the squared norm (variability) of the Gaussian vector after projection onto orthogonal subspaces. (e.g., sum of squares between groups or residuals in ANOVA)
- $\sum_{i=1}^n Z_i^2 = \sum_{k=1}^s Q_k$: decomposition of the total sum of squares into orthogonal components. (cochran assumes that $\sum_{k=1}^s A_k = I_n$ and we have $\sum_{i=1}^n Z_i^2 = Z^T Z = Z^T I_n Z = Z^T[\sum A_k]Z = \sum Q_k$)
- rank$(A_k) = r_k$: the degrees of freedom of each quadratic form.
- Condition $\sum_{k=1}^s r_k = n$: the sum of the degrees of freedom equals the dimension of A, which ensures independence of the quadratic forms.
- Conclusion $Q_k/\sigma^2 \sim \chi^2_{r_k}$: each quadratic form follows a chi-squared distribution with r_k degrees of freedom.

Lemmas:

Lemma 1 (Distribution of the quadratic form of the of a Gaussian vector)**.** *if $Z \sim N_n(0, I_n)$ and A is symetric and idemponent of rank k then it's quadratic form satisfies : $Q_k \sim \chi^2_{r_k}$*

1. *Use the Spectral theorem for symetric matrixes: we have $A = PD^TP$*
   - *P is the matrix of eigenvectors of A ($P^TP = PP^T = I$)*
   - *D is a diagonal matrix of eigenvalues : $D = diag(\lambda_1, \dots, \lambda_n)$*
2. *Since A is idemponent ($A^2 = A$) so $D^2 = D$ this means*

$$\lambda_i^2 = \lambda_i \implies \lambda_i \in \{0, 1\}.$$

   *The number of eigenvalues equal to 1 is the **rank** r of A.*

   The number of eigenvalues equal to 0 is $n - r$.

   Hence, the spectral decomposition of A is

$$D = diag(\underbrace{1, \dots, 1}_{r\ ones}, \underbrace{0, \dots, 0}_{n-r\ zeros}),$$

3. *Change the variables using the orthogonale matrix P , because multiplying a multivariate normal vector by an orthogonal matrix preserves the Gaussian distribution and the covariance*

$$Y = P^TZ$$

4. *rewrite the quadratic form in terms of Y we will find:*

$$Q_k = Z^TAZ = Y^TDY$$

   and because D is diagonal we find:

$$Y^TDY = \sum_{i=1}^{n} \lambda_i Y_i^2$$

   and from the point 2. we find that:

$$Q_k = \sum_{i=1}^{r} Y_i^2$$

   We have:

$$Y_i \sim N(0, 1) \implies Y_i^2 \sim \chi^2_1$$

   Since Y_i are independat, so:

$$Q = \sum_{i=1}^{r} Y_i^2 \sim \chi^2_r$$

Lemma 2 (indépendance of orthogonal quadratic forms)**.** *this lemma says that Q and $Q\prime$ are independent*

1. *$Z = AZ + (I - A)Z$ is an orthogonal decomposition*
2. *if $Q = \sum \lambda_i Y_i^2$ so: $Q = \sum_{i=1}^{r} Y_i^2 \sim \chi^2_r$*
3. *$I - A = P^T(A - D)P$*
4. *eigenvalues of $(I - A)$ are r zeros and $n - r$ ones*
5. *$Q\prime = \sum_{i=r+1}^{n} Y_i^2 \sim \chi^2_{n-r}$*

Finally:

- *Q is a finction of $Y_1 \dots Y_r$*
- *$Q\prime$ is a function of $Y_{r+1} \dots Y_n$*
- *Since these depend on disjoint sets of independent Gaussian variables, so Q and $Q\prime$ are indepndent*

The proof of the theorem:

Since in our theorem's statement we have an **If and only if**, we need to prouve it in both directions:
we assume: $\sigma^2 = 1$

1. $\Longleftarrow$

   - from Lemma1 and Lemma2 Q_k are independent, and $Q_k \sim \chi^2_{r_k}$
   - $\sum_{i=1}^{m} Z_i^2 \sim \chi^2_n$
   - $\sum Q_k \sim \chi^2_{\sum r_k}$ because Q_k are independent, and each folowsa chi squared distribution

2. $\Longrightarrow$

   - assume that $\sum r_k = n$
   - consider $Q = Z^T A_1 Z$ and $A_{-1} = \sum_{k=2}^{n} A_k$ and $A_{-1} = I_n - A_1$
   - so A_1 and A_{-1} are simultaneously diagonalizable (same eigenvalues)
   - $A_1 = P^T D P$ and $A_{-1} = P^T D_{-1} P$
   - $A_1 + A_{-1} = I_n$ so $D_1 + D_{-1} = I_n$. Also: $rank(D_1) = r$ and $rank(D_{-1}) = n - r$
   - There is only one way that this could happen (up to ordering):

   $$D_1 = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \quad D_{-1} = \begin{pmatrix} 0 & 0 \\ 0 & I_{n-r} \end{pmatrix}$$

   Finally: from Lemma2 we have Q_1 and Q_{-1} are independent and follow a chi-squared distribution with a DOF : r_k and $n - r_k$
   andreapiting this argument recursivelyfor $A_2 \dots A_s$ we conclude $\sum_{k=1}^{s} r_k = n$

What we did in this proof, is just for one A_k, the proof do it for the remaining ones $Q_2 \dots Q_S$, and we will end up with s independent quadratic forms that follows chi-squared distribution wew ill take A_2 and we will diagonalized it as before, we will end-up with:

$$D_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & I_{r_2} & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad D_{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I_{n-r_1-r_2} \end{pmatrix}$$

Hence: $Q_1 \sim \chi^2_{r_1}$ and $Q_2 \sim \chi^2_{r_2}$ and $Q_{-2} \sim \chi_{n-r_1-r_2}$ and we will repeat this for $Q_3 \dots Q_s$ to conclude.

One way ANOVA, DoF and Cochran's theorem:

Cochran's theorem explains why the sums of squares in one-way ANOVA are independent and why their degrees of freedom correspond to the dimensions of the associated subspaces.

In one way ANOVA we decompose the variability as follows:

$$Total - SS = Between - groups - SS + Within - groups - SS$$

Cochran's theorem gives the theoretical justification for this decomposition by showing that:

- Each sum of squares is a quadratic form of a Gaussian vector
- These quadratic forms correspond to orthogonal projections
- Each sum of squares follows a χ^2 distribution
- The degrees of freedom are exactly the ranks of the projection matrices
- The sum of the degrees of freedom equals the total sample size minus constraints

So Cochran's theorem explains:

- why the degrees of freedom are what they are
- why sums of squares are independent
- why F-statistics in ANOVA follow an F distribution

Geometrically: The data vector lies in an n-dimensional space, and Cochran's theorem decomposes it into orthogonal linear subspaces; each quadratic form measures the squared length of the projection onto one of these subspaces.