Osborne's Algorithm

In this writeup I will provide the Osborne's algorithm in details. We do not assume any specific knowledge in Statistics. Only some basic calculus and algebra knowledge should be sufficient to understand this notes. We have the following problem at hand. We have the following observations $\{y_1, \ldots, y_n\}$, from the following model

$$y_t = \alpha_1 e^{\beta_1 t} + \alpha_2 e^{\beta_2 t} + e_t; \qquad t = 1, \ldots, n. \tag{1}$$

It is assumed that e_t's have mean zero and equal variances, and we want to estimate α_1, β_1, α_2 and β_2 from $\{y_1, \ldots, y_n\}$. Before progressing further we would like to make few comments. First of all the method we are going to describe is not very specific for only two component models, it can be easily implemented for any arbitrary p component model also as the following

$$y_t = \sum_{k=1}^{p} \alpha_k e^{\beta_k t} + e_t; \qquad t = 1, \ldots, n. \tag{2}$$

But we will discuss about the model (1) for brevity. Another observation is very important. Note that if β_1 and β_2 are known, then α_1 and α_2 can be estimated as the least squares estimators as

$$\begin{bmatrix} \widehat{\alpha}_1 \\ \widehat{\alpha}_2 \end{bmatrix} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y}, \tag{3}$$

where $\mathbf{X}$ is a $n \times 2$ matrix, and $\mathbf{Y}$ is a $n \times 1$ vector as follows

$$\mathbf{X} = \begin{bmatrix} e^{\beta_1} & e^{\beta_2} \\ e^{2\beta_1} & e^{2\beta_2} \\ \ldots & \ldots \\ e^{n\beta_1} & e^{n\beta_2} \end{bmatrix} \quad \text{and} \quad \mathbf{Y} = (y_1, \ldots, y_n)^\top. \tag{4}$$

Further, we will show that if

$$\mu_t = \alpha_1 e^{\beta_1 t} + \alpha_2 e^{\beta_2 t}; \qquad t = 1, \ldots, n, \tag{5}$$

and we know $\mu_1, \ldots, \mu_n$, then how α_1, β_1, α_2 and β_2 can be obtained. First we will show that for a given $\{\mu_1, \ldots, \mu_n\}$, there exists g_0, g_1 and g_2, such that

$$g_0 \mu_k + g_1 \mu_{k+1} + g_2 \mu_{k+2} = 0; \qquad k = 1, \ldots, n-2. \tag{6}$$

Consider the following quadratic equation:

$$(x - e^{\beta_1})(x - e^{\beta_2}) = x^2 - (e^{\beta_1} + e^{\beta_2})x + e^{\beta_1+\beta_2} = 0. \tag{7}$$

The above quadratic equation (7) has roots e^{β_1} and e^{β_2}. Suppose $g_2 = 1$, $g_1 = -(e^{\beta_1} + e^{\beta_2})$ and $g_0 = e^{\beta_1+\beta_2}$, and consider the following expression

$$g_2 e^{\beta_1(k+2)} + g_1 e^{\beta_1(k+1)} + g_0 e^{\beta_1 k} = e^{\beta_1 k}(g_2 e^{2\beta_1} + g_1 e^{\beta_1} + g_0) = 0,$$

because of (7). Similarly,

$$g_2 e^{\beta_2(k+2)} + g_1 e^{\beta_2(k+1)} + g_0 e^{\beta_2 k} = e^{\beta_2 k}(g_2 e^{2\beta_2} + g_1 e^{\beta_2} + g_0) = 0,$$

Hence, (6) also follows for all α_1 and α_2. From (6) we can write that there exists matrices $\mathbf{G}$ of order $(n-2) \times n$ and $\mathbf{M}$ of order $(n-2) \times 3$, such that

$$\mathbf{G}^\top \mathbf{M} = 0;$$

where

$$\mathbf{G}^\top = \begin{bmatrix} g_0 & g_1 & g_2 & 0 & 0 & \dots & \dots \\ 0 & g_0 & g_1 & g_2 & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & g_0 & g_1 & g_2 \end{bmatrix} \quad \text{and} \quad \mathbf{M} = \begin{bmatrix} \mu_1 & \mu_2 & \mu_3 \\ \mu_2 & \mu_3 & \mu_4 \\ \vdots & \vdots & \vdots \\ \mu_{n-2} & \mu_{n-1} & \mu_n \end{bmatrix}.$$

Now we are in a position to provide Osborne's algorithm to compute the least squares estimators of the unknown parameters. If we define $\boldsymbol{\theta} = (\alpha_1, \alpha_2, \beta_1, \beta_2)$, then the LSEs of the unknown parameters can be obtained by minimizing

$$Q(\boldsymbol{\theta}) = \sum_{i=1}^{n} \left(y_t - \alpha_1 e^{\beta_1 t} - \alpha_2 e^{\beta_2 t}\right)^2. \tag{8}$$

Note that (8) can be written as

$$Q(\boldsymbol{\theta}) = (\mathbf{Y} - \mathbf{X}\boldsymbol{\alpha})^\top (\mathbf{Y} - \mathbf{X}\boldsymbol{\alpha}). \tag{9}$$

Here the matrix $\mathbf{X}$ is same as defined before, and it is a function of $\boldsymbol{\beta} = (\beta_1, \beta_2)$, but we do not make it explicit. Further, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)^\top$. Therefore, the minimization of $Q(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$, is equivalent to minimization of $R(\boldsymbol{\beta})$, where

$$\begin{aligned} R(\boldsymbol{\beta}) &= (\mathbf{Y} - \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{Y})^\top(\mathbf{Y} - \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{Y}) \\ &= \mathbf{Y}^\top(\mathbf{I} - \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top)^\top(\mathbf{I} - \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top)\mathbf{Y} \\ &= \mathbf{Y}^\top(\mathbf{I} - \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top)\mathbf{Y} \\ &= \mathbf{Y}^\top(\mathbf{I} - \mathbf{P}_{\mathbf{X}})\mathbf{Y}. \end{aligned}$$

Here

$$\mathbf{P}_{\mathbf{X}} = \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top,$$

is known as the projection operator on the column space of the matrix $\mathbf{X}$. Note that the rank of the matrix $\mathbf{X}$ is 2. Since $\mathbf{G}^\top\mathbf{X} = \mathbf{0}$, and the rank of the matrix $\mathbf{G}$ is $n-2$, hence

$$\mathbf{P}_{\mathbf{X}} + \mathbf{P}_{\mathbf{G}} = \mathbf{I},$$

where

$$\mathbf{P}_{\mathbf{G}} = \mathbf{G}(\mathbf{G}^\top\mathbf{G})^{-1}\mathbf{G}^\top,$$

the projection operator on the column space of the matrix $\mathbf{G}$. The following observation is useful: for every β_1, β_2, there exists one $\mathbf{g} = (g_0, g_1, g_2)^\top$, such that, e^{β_1} and e^{β_2} are the roots of

$$g_2 x^2 + g_1 x + g_0 = 0.$$

Hence, we minimize

$$R(\mathbf{g}) = \mathbf{Y}^\top\mathbf{P}_{\mathbf{G}}\mathbf{Y}$$

with respect to $\mathbf{g}$. The following observations are useful. First note that

$$\mathbf{G} = g_0\mathbf{G}_0 + g_1\mathbf{G}_1 + g_2\mathbf{G}_2,$$

where

$$\mathbf{G}_0^\top = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{G}_1^\top = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & 1 & 0 \end{bmatrix}$$

and

$$\mathbf{G}_2^\top = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 & 0 & 1 \end{bmatrix}.$$

Further,

$$\frac{d}{dg_k}\mathbf{G} = \mathbf{G}_k; \quad k = 0, 1, 2,$$

$$\frac{d}{dg_k}\left\{(\mathbf{G}^\top\mathbf{G})^{-1}\right\} = (\mathbf{G}^\top\mathbf{G})^{-1}\left(\mathbf{G}_k^\top\mathbf{G} + \mathbf{G}^\top\mathbf{G}_k\right)(\mathbf{G}^\top\mathbf{G})^{-1}.$$

and

$$\begin{aligned} \frac{d}{dg_k}\{\mathbf{P}_\mathbf{G}\} &= \frac{d}{dg_k}\left\{\mathbf{G}(\mathbf{G}^\top\mathbf{G})^{-1}\mathbf{G}^\top\right\} \\ &= \mathbf{G}_k(\mathbf{G}^\top\mathbf{G})^{-1}\mathbf{G}^\top + \mathbf{G}(\mathbf{G}^\top\mathbf{G})^{-1}\left(\mathbf{G}_k^\top\mathbf{G} + \mathbf{G}^\top\mathbf{G}_k\right)(\mathbf{G}^\top\mathbf{G})^{-1}\mathbf{G}^\top + \\ &\quad \mathbf{G}(\mathbf{G}^\top\mathbf{G})^{-1}\mathbf{G}_k^\top \end{aligned}$$

Now the three normal equations are

$$\frac{d}{dg_k}R(\mathbf{g}) = 0; \quad k = 0, 1, 2,$$

and they can be written as

$$\mathbf{Y}^\top\left\{\mathbf{G}_k(\mathbf{G}^\top\mathbf{G})^{-1}\mathbf{G}^\top + \mathbf{G}(\mathbf{G}^\top\mathbf{G})^{-1}\left(\mathbf{G}_k^\top\mathbf{G} + \mathbf{G}^\top\mathbf{G}_k\right)(\mathbf{G}^\top\mathbf{G})^{-1}\mathbf{G}^\top + \mathbf{G}(\mathbf{G}^\top\mathbf{G})^{-1}\mathbf{G}_k^\top\right\}\mathbf{Y} = 0$$

for $k = 0, 1, 2$. The above three equations can be written in a matrix form as

$$\mathbf{B}(\mathbf{g})\mathbf{g} = \mathbf{0} \tag{10}$$

here $\mathbf{B}(\mathbf{g})$ is a 3×3 matrix whose (j,k)-th element is b_{jk}, where for $0 \le j, k \le 2$,

$$\begin{aligned} b_{jk} &= \mathbf{Y}^\top\mathbf{G}_k(\mathbf{G}^\top\mathbf{G})^{-1}\mathbf{G}_j^\top\mathbf{Y} + \mathbf{Y}^\top\mathbf{G}(\mathbf{G}^\top\mathbf{G})^{-1}\left(\mathbf{G}_k^\top\mathbf{G}_j + \mathbf{G}_j^\top\mathbf{G}_k\right)(\mathbf{G}^\top\mathbf{G})^{-1}\mathbf{G}^\top\mathbf{Y} \\ &\quad + \mathbf{Y}^\top\mathbf{G}_j(\mathbf{G}^\top\mathbf{G})^{-1}\mathbf{G}_k^\top\mathbf{Y}. \end{aligned}$$

It is immediate that the matrix $\mathbf{B}(\mathbf{g})$ is a symmetric matrix as $b_{jk} = b_{kj}$, hence all the eigenvalues of $\mathbf{B}(\mathbf{g})$ will be real. The solution of (10) is the eigenvector vector $\mathbf{g}$ corresponding to the zero eigenvalue of $\mathbf{B}(\mathbf{g})$.

Hence, we can use the following algorithm to solve the set of normal equations (10).

Algorithm:

Step 1: Start with some initial guess of $\mathbf{g}$, say $\mathbf{g}^{(0)}$. Normalize it, i.e. make it $|\mathbf{g}^{(0)}| = 1$.

Step 2: Compute the eigenvector corresponds to the minimum eigenvalue (in absolute sense) of the matrix $\mathbf{B}(\mathbf{g}^{(0)})$. Suppose it is $\mathbf{g}^{(1)}$.

Step 3: Normalize $\mathbf{g}^{(1)}$, and go back to Step 2. Continue the process until convergence takes place.

Once $\widehat{\mathbf{g}} = (\widehat{g}_0, \widehat{g}_1, \widehat{g}_2)^\top$ is obtained, then find the roots of the quadratic equation

$$\widehat{g}_2 x^2 + \widehat{g}_1 x + \widehat{g}_0 = 0,$$

as $e^{\widehat{\beta}_1}$ and $e^{\widehat{\beta}_2}$. Once $\widehat{\beta}_1$ and $\widehat{\beta}_2$ are obtained, $\widehat{\alpha}_1$ and $\widehat{\alpha}_2$ can be obtained as described before.