پس در این ساختار تنها به برآورد دو پارامتر نیاز داریم و ساختار ماتریس واریانس کوواریانس مشخص تری باشد.

* دقت داشته باشید که این ساختار ماتریس واریانس کوواریانس با فرضیاتی است که ما در نظر گرفتیم (بعنوان مثال وابستگی درون گروهی داریم.) پس باید داده ها را بررسی شود تا بفهمیم که آیا این ساختار برای داده ها درست است یا خیر و مسأله misespicefiation بوجود نیاید. چون باعث می شود برآوردهای OLS ، BLUE نشوند. بعنوان مثال این مدل برای داده های پانلی خوب است ولی برای داده های طولی که وابستگی پیاپی دارند مناسب نمی باشد.

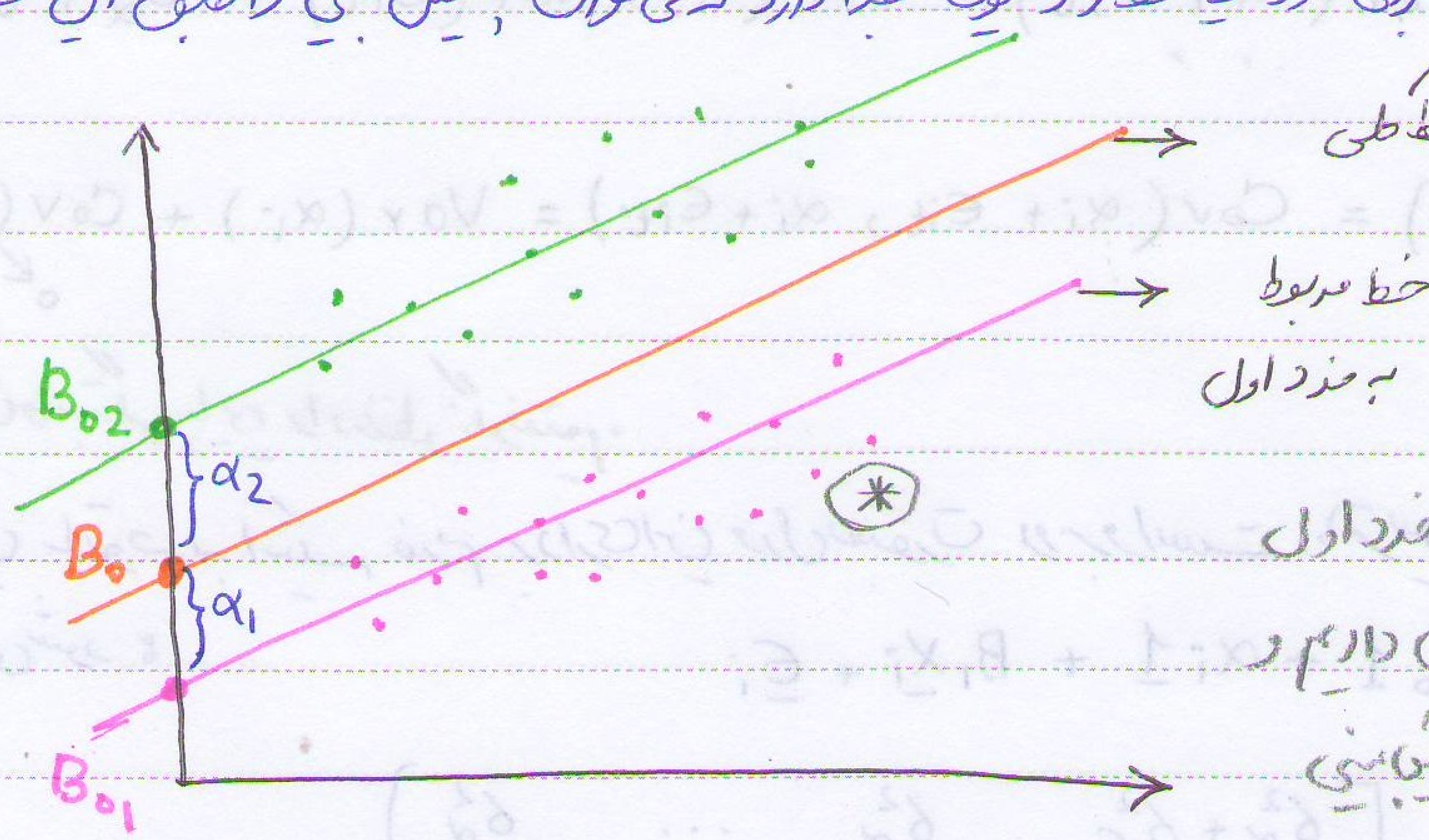
به این مدل، random intercept model می گویند و به صورت زیر آن را نمایش می دهند:

$$y_{it} = \underbrace{B_{0i}}_{B_0 + \alpha_i} + B_1 x_{it} + \epsilon_{it}$$

B_0 ← عرض از مبدأ کلی

α_i → مقدار اختلاف با عرض از مبدأ کلی برای فرد i

در این حالت هر فرد برای خود یک خط رگرسیونی جدا دارد که می توان پیش بینی را طبق این خط برای آن فرد در نظر گرفت:



(*) با استفاده از مشاهدات فرد اول خط رگرسیونی برازش داده ایم و می توانیم، برای این فرد پیش بینی صورت دهیم.

خطوط رگرسیونی برای افراد مختلف دارای شیب یکسان و عرض از مبدأ متفاوت می باشند.

* مقدار $\hat{\alpha}_i$ هم مانند $\hat{\epsilon}_i$ برآورد می‌شود، $\hat{B}_0$ و $\hat{B}_1$ را هم بدست می‌آوریم، پس به این صورت مقدار عرض از مبدأ برای هر فرد برآورد می‌شود و می‌توان خط رگرسیونی مربوط به آن فرد را رسم کرد.

* با استفاده از خط کلی می‌توان برای فرد جدید اضافه شده به نمونه، پیش بینی صورت داد.

2- حالت دوم:

این بار علاوه بر α_i، $b_i x_{it}$ (که اثر تصادفی است) را نیز وارد مدل می‌کنیم. در واقع اینبار علاوه بر عرض از مبدأ، شیب مدل هم تصادفی می‌شود.

$$y_{it} = B_{0i} + B_{1i}\, x_{it} + \epsilon'_{it}$$

$$B_{0i} = \underbrace{B_0}_{\text{عرض از مبدأ کلی مدل}} + \alpha_i \qquad B_{1i} = \underbrace{B_1}_{\text{شیب کلی مدل}} + b_i$$

$$\left\{\begin{array}{l} \underline{\epsilon}'_i \overset{iid}{\sim} N(\underline{0}, \sigma^2_{\epsilon'} I) \\ \epsilon'_{it} \overset{iid}{\sim} N(0, \sigma^2_{\epsilon'}) \\ \text{ind} \\ \begin{pmatrix}\alpha_i \\ b_i\end{pmatrix} \sim N_2(0, \Sigma) \; (*) \end{array}\right.$$

$$\Sigma = \begin{bmatrix} \sigma^2_\alpha & \sigma_{\alpha b} \\ \sigma_{\alpha b} & \sigma^2_b \end{bmatrix}$$

$$y_{it} = B_0 + B_1 x_{it} + \underbrace{\alpha_i + b_i x_{it} + \epsilon'_{it}}_{u_{it}}$$

که قبلاً گفتیم ناهمبسته اند.

(*) α_i و b_i بطور کلی هم وابسته اند، چون عوامل مربوط به فرد i ام هستند، بنابراین بینشان کوواریانس هست و توزیع را توأماً اینجوری در نظر می‌گیریم.

$$\underline{y}_i = [1 \quad x_i]\begin{bmatrix}B_0 \\ B_1\end{bmatrix} + [1 \quad x_i]\begin{bmatrix}\alpha_i \\ b_i\end{bmatrix} + \underline{\epsilon}'_i$$

دقت کنید که

همانطور که از ابتدا گفتیم یک متغیر توضیحی در نظر گرفتیم.

$$\underline{y}_i = Z_i \underline{\beta} + Z_i \underline{\delta}_i + \underline{\epsilon}'_i$$

حال ساختار ماتریس واریانس کواریانس را بدست می آوریم:

$$Var(\underline{y}_i) = Var(z_i \underline{\delta}_i + \underline{\epsilon}'_i) = z_i \, Var(\underline{\delta}_i) \, z_i' + Var(\underline{\epsilon}'_i) =$$

$$= z_i \Sigma z_i' + \sigma^2_{\epsilon'} I$$

البته اگر بخواهیم درایه های ماتریس را بصورت دقیق بدست آوریم، داریم:

$$Var(y_{it}) = Var(\alpha_i + b_i x_{it} + \epsilon'_{it}) = Var(\alpha_i) + x_{it}^2 \, Var(b_i)$$

$$+ Var(\epsilon'_{it}) + 2x_{it} \, Cov(\alpha_i, b_i) = \sigma^2_\alpha + x_{it}^2 \sigma^2_b + \sigma^2_{\epsilon'} + 2x_{it} \sigma_{\alpha b}$$

$$Cov(y_{it}, y_{it'}) = Cov(\alpha_i + b_i x_{it} + \epsilon'_{it}, \; \alpha_i + b_i x_{it'} + \epsilon'_{it'}) =$$

چون وابستگی بینشان نداریم

$$= Var(\alpha_i) + x_{it} x_{it'} \, Var(b_i) + \underbrace{Cov(\epsilon'_{it'}, \epsilon'_{it})}_{0} + x_{it'} \, Cov(\alpha_i, b_i)$$

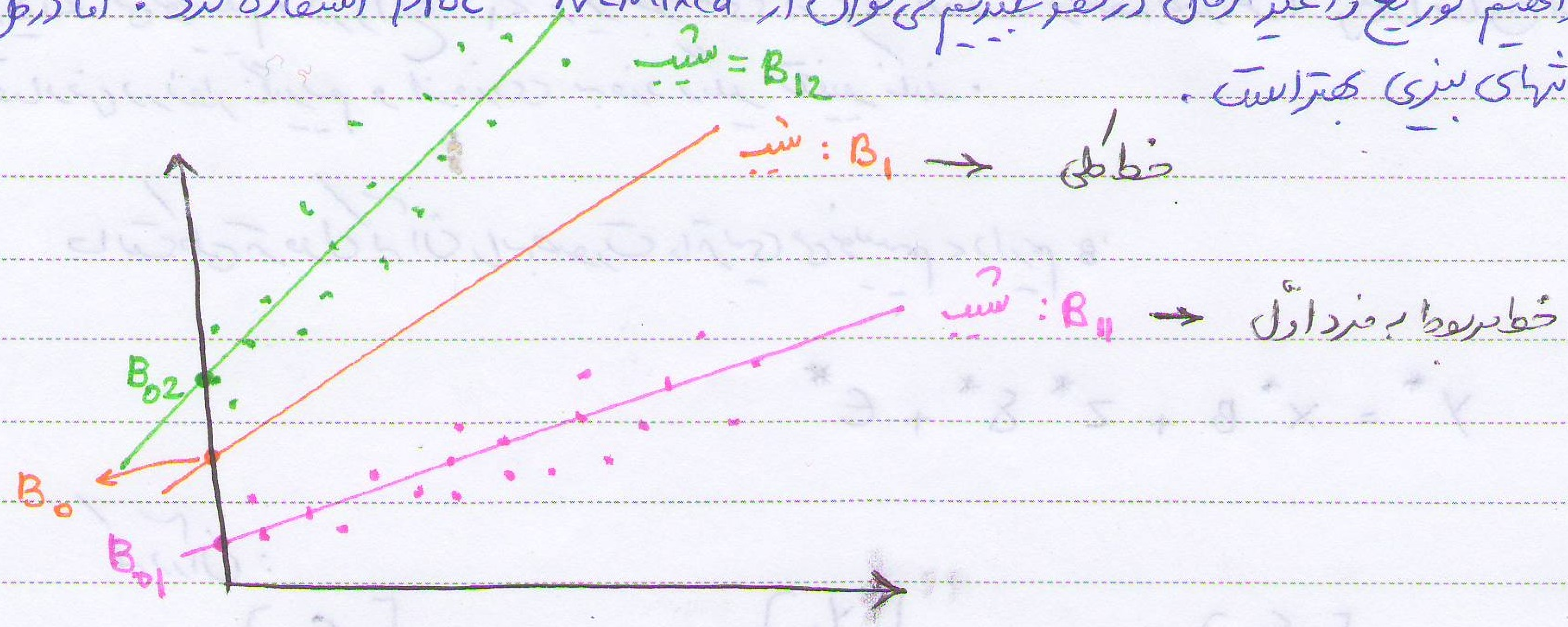
$$+ \underbrace{Cov(\alpha_i, \epsilon'_{it'})}_{0} + x_{it} \, Cov(\alpha_i, b_i) + x_{it} \, \underbrace{Cov(b_i, \epsilon'_{it'})}_{0} + \underbrace{Cov(\epsilon'_{it}, \alpha_i)}_{0}$$

$$+ x_{it'} \, \underbrace{Cov(b_i, \epsilon'_{it})}_{0} = \sigma^2_\alpha + x_{it} x_{it'} \sigma^2_b + x_{it'} \sigma_{\alpha b} + x_{it} \sigma_{\alpha b}$$

$$Var(\underline{y}_i) = \begin{bmatrix} \sigma^2_\alpha + x_{it}^2 \sigma^2_b + \sigma^2_{\epsilon'} + 2x_{it}\sigma_{\alpha b} & \sigma^2_\alpha + x_{it} x_{it'} \sigma^2_b + x_{it'} \sigma_{\alpha b} + x_{it} \sigma_{\alpha b} & \cdots \\ & \sigma^2_\alpha + x_{it}^2 \sigma^2_b + \sigma^2_{\epsilon'} + 2x_{it}\sigma_{b\alpha} & \cdots \\ & & \ddots \end{bmatrix}$$

این ساختار هم چهار پارامتر مجهول دارد که به راحتی قابل برآورد هستند.

در نرم افزار SAS، با استفاده از proc mixed می توان انجام داد، همچنین اگر بخواهیم توزیع را غیر نرمال در نظر بگیریم می توان از proc NLmixed استفاده کرد. امّا در کل روشهای بیزی بهتر است.



این جا نیز برای هر فرد، خط رگرسیونی جدا داریم، فقط در این حالت شیب خطوط نیز با هم فرق دارد.

* در بررسی های طولی تعداد n باید زیاد باشد، چون هدف همان خط کلی با B_0 و B_1 است، پس باید n زیاد باشد. از طرفی برعکس داده های سری زمانی، معمولاً T کم است، چون اگر T زیاد باشد در بررسی ها تغییرات فصلی و ... هم ایجاد می شود.

* حالت کلی: در حالت کلی ممکن است مدل دارای p متغیر توضیحی باشد، بنابراین داریم:

$$\underline{Y_i} = X_i \underline{\beta} + Z_i \delta_i + \epsilon_i \qquad \epsilon_i \overset{iid}{\sim} N_T(0, \sigma^2 I_T)$$

$$\delta_i \sim N_q(\underset{\sim}{0}, \Psi)$$

$$X_i = \begin{bmatrix} 1 & x_{i11} & \cdots & x_{i1p} \\ 1 & x_{i21} & & x_{i2p} \\ \vdots & \vdots & & \vdots \\ 1 & x_{iT1} & & x_{iTp} \end{bmatrix}$$

$\underline{X}_{i1}$ بردار

$\underline{X}_{ip}$ بردار

شامل T مشاهده فرد i که مربوط به متغیر X_1 است.

نکته: دقت کنید که چون ممکن است همه متغیرهای توضیحی را دارای ضرایب تصادفی در نظر نگیریم، زکه Z_i می‌نویسیم. در واقع Z_i زیرمجموعه‌ای از X_i است و همه ضرایب B_j را ممکن است تصادفی در نظر نگیریم و از فردی به فرد دیگر تغییر نکند.

حالت کلی‌تر مدل که آن را به صورت ماتریسی می‌نویسیم، داریم:

$$Y^* = X^* B + Z^* \delta^* + \epsilon^*$$

که در آن:

$$\delta^* = \begin{bmatrix} \underset{\sim}{\delta_1} \\ \underset{\sim}{\delta_2} \\ \vdots \\ \underset{\sim}{\delta_n} \end{bmatrix} , \quad Y^* = \begin{bmatrix} \underset{\sim}{Y_1} \\ \underset{\sim}{Y_2} \\ \vdots \\ \underset{\sim}{Y_n} \end{bmatrix} , \quad \epsilon^* = \begin{bmatrix} \underset{\sim}{\epsilon_1} \\ \underset{\sim}{\epsilon_2} \\ \vdots \\ \underset{\sim}{\epsilon_n} \end{bmatrix}$$

$$X^* = \begin{bmatrix} \underset{\sim}{X_1} & \underset{\sim}{0} & \cdots & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{X_2} & \cdots & \underset{\sim}{0} \\ & & \ddots & \\ & & & \underset{\sim}{X_n} \end{bmatrix} \xrightarrow{} \text{ماتریس} , \quad Z^* = \begin{bmatrix} \underset{\sim}{Z_1} & \underset{\sim}{0} & \cdots & \underset{\sim}{0} \\ & \underset{\sim}{Z_2} & \cdots & \underset{\sim}{0} \\ & & \ddots & \\ & & & \underset{\sim}{Z_n} \end{bmatrix}$$

$$V = Var(Y^*) = \begin{bmatrix} Var(\underline{Y_1}) & \underset{\sim}{0} & \cdots & \underset{\sim}{0} \\ & Var(\underline{Y_2}) & & \\ & & \ddots & \\ & & & Var(\underline{Y_n}) \end{bmatrix}$$

$$Y^* \sim N_{nT}(X^* B , V)$$

$$\hat{\beta}_{GLS} = (X^{*\prime} V^{-1} X^*) X^{*\prime} V^{-1} Y^*$$

* برای اینکه likelihood بنویسیم، باید V معلوم باشد، پس از REML استفاده می‌کنیم و V را برآورد می‌کنیم و معلوم می‌شود.