



FAKULTETEN FÖR
NATURVETENSKAPER OCH TEKNIK

AVHANDLING PRO GRADU I MATEMATIK

Definitiva gallringsplaner och analys av splitplotförsök

Skribent:
Jonatan Juslin, 40378

Handledare:
Mikael Kurula

2022

Abstrakt

I den här pro gradu-avhandlingen behandlas teorin för statistisk datainsamling samt dataanalys. Inom ramen för datainsamling presenteras både blockade försöksplaner, huvudsakligen splitplotplaner, och en familj av gallringsplaner, som kallas definitiva gallringsplaner.

De olika försöksplanernas statistiska egenskaper studeras, främst i relation till linjär regression. Då den bekanta ordinära minstakvadratmetoden inte kan användas med splitplotutförande, så presenteras den generaliserade minstakvadratmetoden och residual maximum likelihood för skattning av kovariansmatrisen. Eftersom residual maximum likelihood är betydligt mer arbetsdryg än ordinära minstakvadratmetoden så presenteras även ekvivalent estimeringsdesign för vilken ordinära minstakvadratmetoden kan användas i stället för residual maximum likelihood.

Innehåll

1	Inledning	4
2	Ordinär linjär regression	6
2.1	Definitioner	6
2.2	Linjär regression	8
2.2.1	Flerdimensionella normalfördelningen	9
2.2.2	Väntevärdesriktighet	9
2.2.3	Responsens kovarians	10
2.2.4	Aliasing och ortogonalitet	11
3	Generaliserad linjär regression	13
3.1	Maximum likelihood	13
3.2	Generaliserade minstakvadratmetoden	16
3.3	Residual maximum likelihood (REML)	18
3.4	Ekvivalent estimeringsdesign (EED)	19
4	Definitiva gallringsplaner	21

<i>INNEHÅLL</i>	3
5 Ortogonal blocking och splitplotplaner	27
5.1 Blockindelade försöksplaner	27
5.2 Ortogonala block	30
5.3 Splitplotplaner	31
5.3.1 Splitplotmodellen	32
5.4 Egenskaper hos en splitplotplan	33
5.4.1 Kostnad för utförande	33
5.5 Analys av splitplotplaner	34
5.5.1 Responsens kovarians	34
5.5.2 Maximum likelihood för splitplotplaner	35
5.5.3 REML för splitplotplaner	37
5.6 Splitplot och DSD	38
A MATLAB-kod för exempel 5.9.	44

Kapitel 1

Inledning

Trots att vi idag har tillgång till stora mängder data så är data ofta av dålig kvalitet och behöver behandlas innan tolkningar kan göras. I denna avhandling presenteras hur man bör gå tillväga för att samla in och analysera bra data. Planeringen av processen att med minsta tänkbara ansträngning kunna samla in data av bra kvalitet och sedan kunna dra så bra slutsatser som möjligt kallas *försöksplanering*. I praktiken genomförs försöksplanering genom att balansera kostnad av datainsamling mot mängden högkvalitativ data, med hjälp av teorin för matematisk statistik.

Faraway presenterar i boken *Practical Regression and ANOVA using R* [3] de grundläggande begreppen för linjär regression samt tillämpningar av dem. Myers et al. presenterar försöksplanering och hur den kan genomföras i boken *Response surface methodology: process and product optimization using designed experiments* [8]. Matematiken som exempelvis Myers et. al. inte alltid går in på i detalj fås ur böckerna *A Course in Mathematical Statistics* av Roussas [9] och *Variance components* av Searle et al. [10].

Sedan splitplotplaner introducerades i mitten på 1920-talet, se [8], så har konceptet ständigt utvecklats. Letsinger, Myers och Lentner [6] introducerade 1996 två olika typer av försöksplaner med slumpningsrestriktioner, som de kallar korsade och icke korsade, varav korsade fallet är det som här kallas splitplotplaner. Anderson och McLean diskuterar ingående problemet med restriktionsfel i boken *Design of Experiments: A Realistic Approach*, se [2, Kap.5]. Vining, Kowalski och Montgomery [11] presenterar 2005 hur *EED-planer* ger samma parameterestimater för ordinära minstakvadratmetoden och generaliserade minstakvadratmetoden vilket förenklar analysen avsevärt. Jones och Nachtsheim [4] skriver år 2009 omfattande om splitplotplaner och hur dess analys bör göras.

År 2011 inför Jones och Nachtsheim [5] en ny familj av gallringsplaner som de döper till *definitiva gallringsplaner*. Xiao, Lin och Bai [12] fördjupar 2012 matematiska förståelsen för dessa gallringsplaner genom att visa hur de kan byggas med hjälp av konferensmatriser. Lin och Yang [7] presenterar 2015 hur splitplotplaner och definitiva gallringsplaner kan kombineras och hur en EED-plan på så vis kan utvecklas. Lin och Yang begränsar sig till fallet med en EED-plan men i denna avhandling presenteras både fallet då vi har en EED-plan och fallet då vi inte har det.

I kapitel 2 introduceras den ordinära minstakvadratmetoden samt flerdimensionella normalfördelningen och några av dess egenskaper. I kapitel 3 diskuteras generaliserade minstakvadratmetoden samt maximum likelihood metoden och när generaliserade minstakvadratmetoden måste användas i stället för ordinära minstakvadratmetoden. Vidare i kapitel 4 diskuteras en relativt ny familj av försöksplaner för gallring av faktorer, kallade definitiva gallringsplaner.

I kapitel 5 diskuteras hur man kan införa olika typer av blocking i sin försöksplan. Både deterministiska och stokastiska blockeffekter diskuteras. Först diskuteras blockindelade försöksplaner i allmänhet varefter en intressant typ av blocking med stokastiska blockeffekter, kallad splitplotplaner, introduceras. Analysen av splitplotplaner diskuteras i detalj, för vilket residual maximum likelihood rekommenderas i litteraturen. Avhandlingen avslutas med ett belysande exempel på hur man kan kombinera definitiva gallringsplaner med splitplot och hur analysen av motsvarande försök bör göras. I bilaga A finns MATLAB-kod med vilken residual maximum likelihood kan göras i allmänhet.

Kapitel 2

Ordinär linjär regression

När data är insamlat behövs metoder för att analysera det. En stor del av försöksplaneringen går ut på att tänka igenom dataanalysen innan man påbörjar datainsamlingen. Detta görs ofta genom att anta att det finns ett linjärt samband mellan de insamlade faktorerna och responsen. Genom att sedan minimera kvadratiske felet som uppstår när linjära sambandet appliceras till data kan skattningar göras. I detta kapitel kommer *ordinära minstakvadratmetoden* samt flerdimensionella normalfördelningen diskuteras.

2.1 Definitioner

Definition 2.1. *Första ordningens* ordinära regressionsmodell är

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i \\ &= \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \epsilon_i, \quad i = 1, 2, \dots, n, \end{aligned} \tag{2.1}$$

där n är antalet försök och $p = k + 1$ antalet parametrar. I modellen kallas y för *responsen*, x_k för *faktorer*, β_j för *parametrar*, *parametern* β_0 för *interceptet* och ϵ för *feltermen*. Feltermen ϵ_i i försök nummer i antas vara $\epsilon_i \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ och ϵ_i är oberoende av ϵ_j för $i \neq j$.

I matrisform blir modell (2.1), se [8, Kap.2.1]:

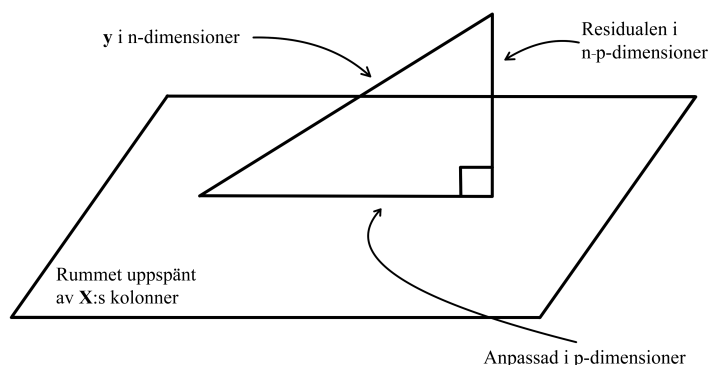
$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}. \quad (2.2)$$

Där $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ är *responsvektorn*, $\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{bmatrix}$ är *datamatriisen*, $\boldsymbol{\beta} =$

$\begin{bmatrix} \beta_0 \\ \vdots \\ \beta_k \end{bmatrix}$ är *parametervektorn* och $\boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix} \in N_n(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ är felet. Här görs antagandet

att $E(\boldsymbol{\epsilon}) = \mathbf{0}$, teorem 2.7 nedan diskuterar mer vad detta antagande och felens fördelning innebär.

Målet med *minstakvadratmetoden* är att kunna skatta den okända $\boldsymbol{\beta}$ med \mathbf{b} i (2.2) så bra som möjligt. Geometriskt sett för n observationer och p parametrar är $\mathbf{y} \in \mathbb{R}^n$ och $\boldsymbol{\beta} \in \mathbb{R}^p$. Den bästa skattningen \mathbf{b} kan geometriskt ges på följande sätt:



Figur 2.1: Geometrisk illustration

Parameterskattningen $\mathbf{b} \approx \boldsymbol{\beta}$ väljs därmed så att felens kvadratsumma minimeras.

$$(\mathbf{y} - \mathbf{X}\mathbf{b})^T(\mathbf{y} - \mathbf{X}\mathbf{b}). \quad (2.3)$$

Då detta expanderas fås:

$$\mathbf{y}^T \mathbf{y} - 2\mathbf{b} \mathbf{X}^T \mathbf{y} + \mathbf{b}^T \mathbf{X}^T \mathbf{X} \mathbf{b}.$$

Genom att derivera partiellt med avseende på alla b_i och sätta derivatorna lika med noll fås vad som kallas *normalekvationen*,

$$\mathbf{X}^T \mathbf{X} \mathbf{b} = \mathbf{X}^T \mathbf{y}.$$

Om *informationsmatrisen* $\mathbf{X}^T \mathbf{X}$ är inverterbar fås slutligen *ordinära minstakvadratstimatet*, se [3, Kap.2.5]

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}. \quad (2.4)$$

Teorem 2.2. *Matrisen $\mathbf{X}^T \mathbf{X}$ är inverterbar om och endast om kolonnerna i \mathbf{X} är linjärt oberoende.*

Bevis. Matrisen $\mathbf{X}^T \mathbf{X}$ är kvadratisk så inverterbar om $N(\mathbf{X}^T \mathbf{X}) = \{0\}$. Vidare är kolonnerna i \mathbf{X} oberoende om och endast om $N(\mathbf{X}) = \{0\}$. Vi bevisar att \mathbf{X} och $\mathbf{X}^T \mathbf{X}$ har samma nollrum, det vill säga $N(\mathbf{X}) = N(\mathbf{X}^T \mathbf{X})$.

Låt $\mathbf{v} \in N(\mathbf{X})$ vilket ger $\mathbf{X} \mathbf{v} = \mathbf{0}$. Då gäller:

$$\mathbf{X}^T \mathbf{X} \mathbf{v} = \mathbf{0} \implies \mathbf{v} \in N(\mathbf{X}^T \mathbf{X}) \implies N(\mathbf{X}) \subset N(\mathbf{X}^T \mathbf{X}).$$

Omvänt, låt $\mathbf{X}^T \mathbf{X} \mathbf{v} = \mathbf{0}$, då gäller:

$$\begin{aligned} \mathbf{v}^T \mathbf{X}^T \mathbf{X} \mathbf{v} &= \mathbf{0} \\ \implies (\mathbf{X} \mathbf{v})^T \mathbf{X} \mathbf{v} &= \mathbf{0} \\ \implies \|\mathbf{X} \mathbf{v}\|^2 &= \mathbf{0} \implies \mathbf{X} \mathbf{v} = \mathbf{0}. \end{aligned}$$

Därmed fås att $\mathbf{v} \in N(\mathbf{X})$ och $N(\mathbf{X}) = N(\mathbf{X}^T \mathbf{X})$. □

2.2 Linjär regression

I detta avsnitt presenteras flerdimensionella normalfördelningen och några av dess grundläggande egenskaper. För mer bakgrund till flerdimensionella normalfördelningen, se [9, Kap. 18.1].

2.2.1 Flerdimensionella normalfördelningen

Definition 2.3. Anta att u_1, u_2, \dots, u_n är oberoende standardnormalfördelade stokastiska variabler och att $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ är vektorn som byggs upp av dem. Då sägs \mathbf{u} vara n -dimensionellt standardnormalfördelat det vill säga $\mathbf{u} \sim N_n(\mathbf{0}_n, \mathbf{I}_n)$.

Definition 2.4. En stokastisk variabel $\mathbf{x} \in \mathbb{R}^n$ sägs vara n -dimensionellt normalfördelat med väntevärdesvektorn $\boldsymbol{\mu} \in \mathbb{R}^n$ och kovariansmatrisen $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$, vilket skrivs $\mathbf{x} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, om \mathbf{x} kan skrivas som $\mathbf{x} = \mathbf{P}\mathbf{u} + \boldsymbol{\mu}$, var $\mathbf{u} \sim N_n(\mathbf{0}_n, \mathbf{I}_n)$ och $\boldsymbol{\Sigma} = \mathbf{P}\mathbf{P}^T$.

Fördelningen är singulär om $\boldsymbol{\Sigma}$ är en singulär matris, annars är fördelningen icke-singulär.

Täthetsfunktionen för en stokastisk variabel som är $\mathbf{x} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, ges av:

$$f(\mathbf{x}) = \frac{e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}}{(2\pi)^{\frac{n}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}}, \quad (2.5)$$

förutsatt att $\boldsymbol{\Sigma}$ är icke-singulär.

2.2.2 Väntevärdesriktighet

Definition 2.5. Punktskattningen $\mathbf{b} \in \mathbb{R}^p$ av parametervektorn $\boldsymbol{\beta} \in \mathbb{R}^p$ sägs vara väntevärdesriktig om:

$$E(\mathbf{b}) = \boldsymbol{\beta} \quad \text{för alla } \boldsymbol{\beta} \in \mathbb{R}^p.$$

Teorem 2.6. Ordinära minstakvadratestimaten för (2.2) är väntevärdesriktigt om $E(\boldsymbol{\epsilon}) = 0$.

Bevis. Låt modellen vara definierad som i (2.2), då fås:

$$\begin{aligned}
E(\mathbf{b}) &= E((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}) \\
&= E((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{X} \boldsymbol{\beta} + \boldsymbol{\epsilon})) \\
&= E((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} \boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\epsilon}) \\
&= E(\boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\epsilon}) \\
&= E(\boldsymbol{\beta}) + E((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\epsilon}) \\
&= \boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T E(\boldsymbol{\epsilon}) \\
&= \boldsymbol{\beta}.
\end{aligned}$$

□

2.2.3 Responsens kovarians

I fortsättningen av detta kapitel antas att felen ϵ_i, ϵ_j är oberoende för $i \neq j$.

Teorem 2.7. *Fördelningen för \mathbf{y} i (2.2) är $\mathbf{y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I})$.*

Bevis. Låt modell (2.2) vara given. Definition 2.4 ger påståendet gällande normalfördelningen. I definition 2.1 antogs $\epsilon_i \sim N(0, \sigma^2 \mathbf{I})$ vilket ger att

$$E(\mathbf{y}) = E(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}) = \mathbf{X}\boldsymbol{\beta}$$

och

$$\begin{aligned}
\text{Var}(\mathbf{y}) &= \text{Var}(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}) = \text{Var}(\mathbf{X}\boldsymbol{\beta}) + \text{Var}(\boldsymbol{\epsilon}) + 2\text{Cov}(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\epsilon}) \\
&= \text{Var}(\boldsymbol{\epsilon}) = \sigma^2 \mathbf{I}.
\end{aligned}$$

□

Teorem 2.8. *Punktskattningen \mathbf{b} av parametervektorn $\boldsymbol{\beta}$ har fördelningen*

$$\mathbf{b} \sim N_p(\boldsymbol{\beta}, \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}).$$

Bevis. Minstakvadratestimaten för (2.2) ges av (2.4) och definition 2.4 ger påståendet gällande normalfördelning. I teorem 2.6 visades att \mathbf{b} är väntevärdesriktigt så det återstår att visa kovariansen.

med

$$\text{Var}(\mathbf{A}\mathbf{y}) = \mathbf{A}\text{Var}(\mathbf{y})\mathbf{A}^T$$

fås

$$\begin{aligned} \text{Var}(\mathbf{b}) &= \text{Var}((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}) \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \text{Var}(\mathbf{y}) [(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T]^T \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \text{Var}(\mathbf{y}) \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \sigma^2 \mathbf{I} \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \\ &= \sigma^2 \mathbf{I} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \\ &= \sigma^2 \cdot (\mathbf{X}^T \mathbf{X})^{-1}. \end{aligned}$$

□

Definition 2.9. En modell sägs vara *BLUE* om den är den *bästa linjära väntevärdesriktiga punktskattningen*, (eng. *Best Linear Unbiased Estimator*). Här innebär *linjär* att \mathbf{b} är en linjär funktion av \mathbf{y} och med *bästa linjära punktskattning* menas att \mathbf{b} har samma eller mindre kovarians än alla andra linjära estimat av $\boldsymbol{\beta}$.

Teorem 2.10. *Ordinära minstakvadratestimaten \mathbf{b} i (2.4) är BLUE.*

Teorem 2.10 kommer senare i teorem 3.6 visas för allmänna fallet då $\epsilon_i \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ i stället för $\epsilon_i \sim N(0, \sigma^2 \mathbf{I})$ som i det ordinära fallet.

2.2.4 Aliasing och ortogonalitet

Anta att vi är intresserade av att anpassa modellen (2.2) men den korrekta modellen är

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{x}_0\alpha_0 + \boldsymbol{\epsilon}. \quad (2.6)$$

Då vi inte beaktar termen $\mathbf{x}_0\alpha_0$ i vår modell så införs ett systematiskt fel som kan leda till att parameterestimaten inte längre är väntevärdesriktigt. Detta fenomen

kallas för *aliasing* vilket innebär att man tror att man skattar någonting men egentligen skattar man något annat på grund av systematiskt fel som införs av en annan parameter.

Anta att vi anpassar modell (2.2) fast modellen egentligen är (2.6). Då kan $E(\mathbf{b})$ på samma vis som i teorem 2.6 beräknas till:

$$E(\mathbf{b}) = \boldsymbol{\beta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{x}_0 \alpha_0. \quad (2.7)$$

Definition 2.11. Matrisen $(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{x}_0$ kallas för *aliasmatrisen* från α_0 till $\boldsymbol{\beta}$ och betecknas med \mathbf{A} .

Därmed fås att (2.7) kan skrivas som

$$E(\mathbf{b}) = \boldsymbol{\beta} + \mathbf{A} \alpha_0.$$

I praktiken kan vi inte påverka den bortlämnade effekten α_0 i definition 2.11 alls. För att kunna undvika eller minimera aliasing måste därmed försöksplanen göras på ett sådant sätt som minimerar tänkbara aliasmatriser \mathbf{A} .

Ett bra sätt att göra det på är att låta kolonnerna i datamatrisen \mathbf{X} vara ortogonala mot varandra det vill säga matrisen $\mathbf{X}^T \mathbf{X}$ är en diagonalmatrix. Då $\mathbf{X}^T \mathbf{X}$ är en diagonalmatrix innebär att faktorer kan utelämnas utan att aliasing uppstår och därmed fås att parameterestimaten är fria från aliasing. Detta kallas för en *ortogonal försöksplan*.

Kapitel 3

Generaliserad linjär regression

I modellen (2.2) görs antagandet att vi har en konstant varians i feltermen ϵ_i för alla $i = 1, \dots, n$. Med andra ord är $\text{Var}(\epsilon_i) = \sigma^2$. Det är dock inte alltid fallet utan det är fullt möjligt att feltermens varians varierar vilket kallas för *heteroskedasticitet*. Vidare är det möjligt att feltermens varianser inte är oberoende, detta kallas för *autokorrelation*.

Då strukturen hos Σ är annorlunda så krävs även alternativa metoder för analys vilket kommer att diskuteras i detta kapitel.

3.1 Maximum likelihood

Definition 3.1. Låt $R(\mathbf{P})$ beteckna kolonnrummet för matrisen \mathbf{P} och låt $N(\mathbf{P})$ beteckna dess nollrum. Matrisen \mathbf{P} är en *projektion* om $\mathbf{P}^2 = \mathbf{P}$. Projektionen \mathbf{P} *projicerar på underrummet* $V \subset \mathbb{R}^n$ om $V = R(\mathbf{P})$ och projektionen \mathbf{P} sägs vara *ortogonal* om $N(\mathbf{P}) \perp R(\mathbf{P})$.

Idén med *maximum likelihood* är att justera fördelnings parametrar så, att ett givet sampel blir det mest sannolika. Enligt (2.5) så har $\mathbf{y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \Sigma)$ täthetsfunktionen

$$f(\mathbf{y}) = \frac{e^{-\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T \Sigma^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})}}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}}. \quad (3.1)$$

Målet är att för givna \mathbf{X} och \mathbf{y} välja $\boldsymbol{\beta}$ och $\boldsymbol{\Sigma}$ så att $L(\boldsymbol{\beta}, \boldsymbol{\Sigma} | \mathbf{y}) := f(\mathbf{y})$ maximeras. Maximeringen görs enklast genom att maximera logaritmen av L , här kallad ℓ :

$$\ell = \log(L) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}). \quad (3.2)$$

För att maximera (3.2) sökes sedan nollställena till partiella derivatorna med avseende på både $\boldsymbol{\beta}$ och $\boldsymbol{\Sigma}$. För att beräkna partiella derivatorna med avseende på komponenterna i $\boldsymbol{\beta}$ kan vi först utveckla uttrycket $-\frac{1}{2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$

$$\begin{aligned} & -\frac{1}{2} \mathbf{y}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} + \frac{1}{2} \mathbf{y}^T \boldsymbol{\Sigma}^{-1} \mathbf{X}\boldsymbol{\beta} + \frac{1}{2} (\mathbf{X}\boldsymbol{\beta})^T \boldsymbol{\Sigma}^{-1} \mathbf{y} - \frac{1}{2} (\boldsymbol{\beta}^T \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X}\boldsymbol{\beta}) \\ & = -\frac{1}{2} \mathbf{y}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} + (\mathbf{X}\boldsymbol{\beta})^T \boldsymbol{\Sigma}^{-1} \mathbf{y} - \frac{1}{2} (\boldsymbol{\beta}^T \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X}\boldsymbol{\beta}). \end{aligned}$$

Partiella derivatan blir då, med två följande räkneregler, se [10, Appendix M.7],

$$\frac{\partial}{\partial \mathbf{x}} \mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{A} \mathbf{x} + \mathbf{A}^T \mathbf{x},$$

och

$$\frac{\partial}{\partial \mathbf{x}} \mathbf{x}^T \mathbf{a} = \mathbf{a} :$$

$$\begin{aligned} \ell_{\boldsymbol{\beta}} &= \frac{\partial \ell}{\partial \boldsymbol{\beta}} = \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} - \frac{1}{2} [\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X}\boldsymbol{\beta} + (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^T \boldsymbol{\beta}] \\ &= \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} - \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X}\boldsymbol{\beta}. \end{aligned} \quad (3.3)$$

Då nollställena till (3.3) sökes fås vad som kallas *maximum likelihood lösningarna till $\boldsymbol{\beta}$* :

$$\mathbf{X}^T \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{y} = \mathbf{X}^T \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{X}\mathbf{b}. \quad (3.4)$$

Definition 3.2. Maximum likelihood lösningen (3.4) kallas för *möjliga generaliserade minstakvadratskattningen* (eng. *feasible generalized least squares*) av parame-tervektorn $\boldsymbol{\beta}$:

$$\mathbf{b}_{FGLS} = (\mathbf{X}^T \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{y}.$$

Exempel 3.3. Anta att vår modell är $\mathbf{y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ det vill säga $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}_n$, och att vi vill använda maximum likelihood för att skatta $\boldsymbol{\beta}$ med \mathbf{b} och $\sigma^2 > 0$ med $\hat{\sigma}^2$.

Log-likelihood funktionen (3.2) kan då förenklas, eftersom

$$\log|\boldsymbol{\Sigma}| = \log(\sigma^{2n}|\mathbf{I}|) = \log(\sigma^{2n}) = n \cdot \log(\sigma^2),$$

och

$$\boldsymbol{\Sigma}^{-1} = (\sigma^2\mathbf{I})^{-1} = \frac{1}{\sigma^2}\mathbf{I},$$

så

$$-\frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T \boldsymbol{\Sigma}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = -\frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}).$$

Med dessa kombinerade fås därmed att (3.2) kan skrivas som

$$\log(L) = -\frac{n}{2}\log(2\pi) - \frac{n}{2} \cdot \log(\sigma^2) - \frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}).$$

Målet är nu att maximera denna med avseende på både $\boldsymbol{\beta}$ och σ^2 .

I (3.2) beror endast termen $-\frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ av $\boldsymbol{\beta}$. För att maximera uttrycket bör därmed denna term minimeras. I denna term är $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ felens kvadratsumma, se (2.3). Denna vet vi från (2.4) att minimeras då $\boldsymbol{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$. Därmed maximeras $\log(L)$ då $\boldsymbol{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$, oberoende av värdet på $\sigma^2 > 0$.

Givet definition 3.1 fås att ortogonala projektmatriken på $R(\mathbf{X})$ ges av:

$$\mathbf{P} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T.$$

För partiella derivatan av ℓ med avseende på σ fås då att

$$\mathbf{X}(\boldsymbol{\beta} - \mathbf{b}) = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\boldsymbol{\epsilon} = \mathbf{P}\boldsymbol{\epsilon},$$

vilket ger

$$\mathbf{r} = \mathbf{X}(\boldsymbol{\beta} - \mathbf{b}) + \boldsymbol{\epsilon} = (\mathbf{I} - \mathbf{P})\boldsymbol{\epsilon}.$$

Då kan residualens kvadratsumma uttryckas som

$$\mathbf{r}^T\mathbf{r} = \boldsymbol{\epsilon}^T(\mathbf{I} - \mathbf{P})^T(\mathbf{I} - \mathbf{P})\boldsymbol{\epsilon} = \boldsymbol{\epsilon}^T(\mathbf{I} - \mathbf{P})\boldsymbol{\epsilon}.$$

Partiella derivatan av ℓ med avseende på σ blir då

$$\ell_\sigma = \frac{\partial \ell}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \cdot \mathbf{r}^T(\mathbf{I} - \mathbf{P})\mathbf{r} \quad (3.5)$$

Nollstället till (3.5) ges av $\hat{\sigma}$, som uppfyller

$$\begin{aligned} -\frac{n}{\hat{\sigma}} + \frac{1}{\hat{\sigma}^3} \cdot \mathbf{r}^T(\mathbf{I} - \mathbf{P})\mathbf{r} &= 0 \\ \hat{\sigma}^2 &= \frac{\mathbf{r}^T(\mathbf{I} - \mathbf{P})\mathbf{r}}{n}. \end{aligned} \quad (3.6)$$

Problemet med maximum likelihood är att denna skattning av σ^2 är inte väntevärdesriktig, ty

$$E(\sigma^2) = \frac{1}{n-p} \mathbf{r}^T \mathbf{r}, \quad (3.7)$$

se [10, Kap.6.6].

Teorem 3.4. *Maximum likelihood skattningen $\hat{\sigma}^2$ av σ^2 är aldrig väntevärdesriktig.*

Bevis. För att maximum likelihood skattningen $\hat{\sigma}^2$ ska vara väntevärdesriktig så skulle i (3.6) krävas att

$$\frac{\mathbf{I} - \mathbf{P}}{n} = \frac{\mathbf{I}}{n-a} \iff \mathbf{P} = -\frac{a}{n-a} \mathbf{I}.$$

Detta innebär att $\text{ran}(\mathbf{P}) = \text{ran}(\mathbf{I})$ som ger att \mathbf{X} är surjektiv och injektiv, alltså kvadratisk. Alltså är

$$\begin{aligned} \mathbf{P} &= \mathbf{I} \\ \implies n - a &= -a \iff n = 0. \end{aligned}$$

Därmed är maximum likelihood skattningen $\hat{\sigma}^2$ av σ^2 är aldrig väntevärdesriktig. \square

3.2 Generaliserade minstakvadratmetoden

Definition 3.5. Modellen för *generaliserade minstakvadratmetoden* är följande:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad (3.8)$$

där $\boldsymbol{\epsilon} \sim N_n(\mathbf{0}, \boldsymbol{\Sigma})$ och $\boldsymbol{\Sigma}$ är en icke singular, positivt semidefinit matris.

Modellen är för övrigt den samma som i (2.2) med skillnaden att feltermernas kovariansstruktur är mera komplicerad. I (2.2) var $\Sigma = \sigma^2 \mathbf{I}$. Vid heteroskedasticitet är Σ fortfarande en diagonalmatrix men diagonalelementen varierar och vid autokorrelation är Σ inte längre en diagonalmatrix, se [3, Kap. 5.1].

I definition 3.2 gavs möjliga generaliserade minstakvadratskattningen. För det allmänna fallet blir den generaliserade minstakvadratskattningen av parametervektorn β :

$$\mathbf{b}_{GLS} = (\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Sigma^{-1} \mathbf{y}. \quad (3.9)$$

På samma vis som för (2.4) kan visas att estimatoren \mathbf{b}_{GLS} är väntevärdesriktig, beviset följer analogt beviset till teorem 2.6.

Teorem 3.6. *Generaliserade minstakvadratestimaten \mathbf{b}_{GLS} i (3.9) är BLUE.*

Bevis. Påståendet om att \mathbf{b}_{GLS} är en estimator är trivialt, från (3.9) fås att \mathbf{b}_{GLS} är en linjär estimator och väntevärdesriktighet motiverades redan. Därmed återstår att visa att \mathbf{b}_{GLS} är det bästa estimaten med dessa egenskaper. På samma vis som i teorem 2.8 fås att kovariansen för \mathbf{b}_{GLS} är

$$\text{Var}(\mathbf{b}_{GLS}) = (\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1}. \quad (3.10)$$

Låt \mathbf{b}_2 vara en alternativ linjär väntevärdesriktig estimator sådan att $\mathbf{b}_2 = \mathbf{E}\mathbf{y}$,

och definiera

$$\mathbf{F} := \mathbf{E} - (\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Sigma^{-1}.$$

Eftersom \mathbf{b}_2 definierades vara väntevärdesriktigt fås

$$E(\mathbf{b}_2) = E(\mathbf{E}\mathbf{y}) = \mathbf{E}\mathbf{X}\beta = \beta,$$

vilket innebär att $\mathbf{E}\mathbf{X} = \mathbf{I}$ och

$$\begin{aligned} \mathbf{F}\mathbf{X} &= \mathbf{E}\mathbf{X} - (\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Sigma^{-1} \mathbf{X} \\ &= \mathbf{I} - \mathbf{I} = \mathbf{0}. \end{aligned}$$

Kovariansen för \mathbf{b}_2 kan då skrivas som:

$$\begin{aligned} \text{Var}(\mathbf{b}_2) &= [(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\Sigma}^{-1} + \mathbf{F}] \cdot \text{Var}(\mathbf{y}) \cdot [(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\Sigma}^{-1} + \mathbf{F}]^T \\ &= [(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\Sigma}^{-1} + \mathbf{F}] \boldsymbol{\Sigma} [(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\Sigma}^{-1} + \mathbf{F}]^T \\ &= (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} + \mathbf{F} \boldsymbol{\Sigma} \mathbf{F}^T + (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} (\mathbf{F} \mathbf{X})^T + \mathbf{F} \mathbf{X} (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \\ &= \text{Var}(\mathbf{b}_{GLS}) + \mathbf{F} \boldsymbol{\Sigma} \mathbf{F}^T. \end{aligned}$$

Eftersom $\mathbf{F} \boldsymbol{\Sigma} \mathbf{F}^T$ är en positiv semidefinit matris fås att $\text{Var}(\mathbf{b}_2) \geq \text{Var}(\mathbf{b}_{GLS})$ i matrismening. Därmed är \mathbf{b}_{GLS} BLUE. \square

3.3 Residual maximum likelihood (REML)

Då maximum likelihood skattningen $\hat{\sigma}^2$ av σ^2 är inte väntevärdesriktig behövs alternativa metoder för att skatta σ^2 . Lösningen till detta problem är att använda residualen i stället för responsen för att skatta σ^2 . Detta görs genom att transformera om responsvektorn \mathbf{y} så att den har väntevärdet noll, oberoende av vad $\boldsymbol{\beta}$ är. Det vill säga vi väljer en matris \mathbf{M} för att transformera vektorn \mathbf{y} på följande vis:

$$\mathbf{r} = \mathbf{M} \mathbf{y} \quad (3.11)$$

så att

$$E(\mathbf{r}) = \mathbf{M} \mathbf{X} \boldsymbol{\beta} = \mathbf{0} \quad \text{för alla } \boldsymbol{\beta}.$$

För att hitta en matris \mathbf{M} sådan att $\mathbf{M} \mathbf{X} = \mathbf{0}$ används *ortogonal projektionsmatrisen* \mathbf{P} på $R(\mathbf{X})$ i \mathbb{R}^n .

Ortogonal projektionsmatrisen på $R(\mathbf{X})$ ges som tidigare av:

$$\mathbf{P} = \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T.$$

Genom att därmed låta $\mathbf{M} = \mathbf{I} - \mathbf{P}$ fås att:

$$(\mathbf{I} - \mathbf{P}) \mathbf{X} = \mathbf{X} - \mathbf{P} \mathbf{X} = \mathbf{X} - \mathbf{X} = \mathbf{0}.$$

Därmed fås att (3.11) blir $\mathbf{r} = (\mathbf{I} - \mathbf{P}) \mathbf{y} = \mathbf{y} - \mathbf{P} \mathbf{y}$ som är residualen för \mathbf{y} och $\mathbf{r} \sim N(0, \mathbf{M} \boldsymbol{\Sigma} \mathbf{M}^T)$. Då blir motsvarigheten till (3.2) för residual likelihood, här kallad $r\ell$ -funktionen,

$$r\ell = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |\mathbf{M} \boldsymbol{\Sigma} \mathbf{M}^T| - \frac{1}{2} \mathbf{r}^T (\mathbf{M} \boldsymbol{\Sigma} \mathbf{M}^T)^{-1} \mathbf{r}. \quad (3.12)$$

Problemet som uppstår är att $\mathbf{M} = \mathbf{I} - \mathbf{P}$ är inte full rang,

$$\text{rang}(\mathbf{I} - \mathbf{P}) = n - (p + 1),$$

och \mathbf{M} är en $n \times n$ matris. Detta innebär att fördelningen \mathbf{r} är *singulär* vilket betyder att fördelningen av \mathbf{r} är koncentrerad på ett $n - (p + 1)$ underrum av \mathbb{R}^n vilket leder till att likelihoodfunktionen blir odefinierad. Lösningen till detta problem är att ersätta \mathbf{M} med $n - (p + 1)$ linjärt oberoende rader av \mathbf{M} .

$$\mathbf{M}_{n \times n} = \mathbf{I} - \mathbf{P} \longrightarrow \widetilde{\mathbf{M}}_{n-(p+1) \times n}.$$

Därefter används maximum likelihood för $\tilde{\mathbf{r}} = \widetilde{\mathbf{M}}\mathbf{y}$ för att estimeras Σ ur residualen:

$$r\ell = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log\left|\widetilde{\mathbf{M}}\Sigma\widetilde{\mathbf{M}}^T\right| - \frac{1}{2}\tilde{\mathbf{r}}^T(\widetilde{\mathbf{M}}\Sigma\widetilde{\mathbf{M}}^T)^{-1}\tilde{\mathbf{r}}. \quad (3.13)$$

Residual maximum likelihood kommer konkretiseras för splitplotplaner i avsnitt 5.5.3 och exempel 5.9.

3.4 Ekvivalent estimeringsdesign (EED)

Ekvivalenta estimeringsdesigner är en speciell typ av försöksplaner var punktskattning av parametrar ger samma resultat för ordinära och generaliserade minstakvadratmetoden. Detta förenklar analysen mycket eftersom i stället för att vara tvungen att använda sig av exempelvis REML, som beskrivet i avsnitt 3.3, så kan man använda sig av den betydligt enklare ordinära minstakvadratmetoden för att analysera av splitplotplaner.

Definition 3.7. En försöksplan är då EED om det finns en inverterbar matris \mathbf{F} , sådan att:

$$\Sigma\mathbf{X} = \mathbf{X}\mathbf{F}. \quad (3.14)$$

Teorem 3.8. Ordinära minstakvadratestimaten \mathbf{b} ger samma estimat som det generaliserade minstakvadratestimaten \mathbf{b}_{GLS} i en EED-plan och enda kandidaten till matrisen \mathbf{F} är:

$$\mathbf{F} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\Sigma\mathbf{X}. \quad (3.15)$$

Bevis. Då Σ och F är inverterbara fås ur (3.14) att

$$\mathbf{X}\mathbf{F}^{-1} = \Sigma^{-1}\mathbf{X}.$$

Givet (3.9) och (2.4) fås att

$$\begin{aligned} \mathbf{b}_{GLS} &= (\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Sigma^{-1} \mathbf{y} \\ &= ((\mathbf{X}\mathbf{F}^{-1})^T \mathbf{X})^{-1} (\mathbf{X}\mathbf{F}^{-1})^T \mathbf{y} \\ &= (\mathbf{F}^{-T} \mathbf{X}^T \mathbf{X})^{-1} \mathbf{F}^{-T} \mathbf{X}^T \mathbf{y} \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{F}^T \mathbf{F}^{-T} \mathbf{X}^T \mathbf{y} \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \\ &= \mathbf{b}. \end{aligned}$$

Gällande (3.15) har vi, givet (3.14), att

$$\begin{aligned} \Sigma \mathbf{X} = \mathbf{X}\mathbf{F} &\implies \mathbf{X}^T \mathbf{X}\mathbf{F} = \mathbf{X}^T \Sigma \mathbf{X} \\ &\implies \mathbf{F} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \Sigma \mathbf{X}. \end{aligned}$$

□

Korollarium 3.9. Om datamatrixen \mathbf{X} är kvadratisk, så är försöksplanen EED .

Bevis. Då \mathbf{X} också är injektiv, så är \mathbf{X} inverterbar och (3.15) förenklas då till

$$\mathbf{F} = \mathbf{X}^{-1} \Sigma \mathbf{X},$$

som uppfyller (3.14).

□

Kapitel 4

Definitiva gallringsplaner

För att få bra mätdata när man gör ett experiment lönar det sig att satsa tid på att fundera kring vilka faktorinställningar man bör samla in responsdata på, vilket ofta är en väldigt tidskrävande och dyr process. Därför vill man gärna utelämna de faktorer som inte har en stor effekt på den respons som man är intresserad av att undersöka.

Processen att hitta faktorer som saknar signifikant inverkan på responsen kallas för *gallring* och är ofta relevant när man samlar in data. Trots det vill man att ens insamlade data skall kunna estimeras modellen så bra som möjligt.

Anta att vi är intresserade av att anpassa första ordningens regressionsmodell, (2.1). Gallring innebär då att vi vill gallra bort parametrar var $\beta_i \approx 0$ för $i = 1, \dots, k$. Därmed får vi en modell som så bra som möjligt kan estimeras modellen med färre försök än om man körde alla faktorkombinationer.

År 2011 introducerade Jones och Nachtsheim en ny familj av gallringsplaner som rätt bra klarar av att estimeras första ordningens modeller samt en del interaktionstermer med få försök. Dessa gallringsplaner kallas för *definitiva gallringsplaner*, (eng. *Definitive Screening Designs* eller *DSD*), se [5].

Den huvudsakliga fördelen med dessa planer är att de kräver betydligt färre försök än exempelvis de traditionella fullständiga 2^k planerna, se [8, Kap.3.4]. För m faktorer kräver definitiva gallringsplaner endast $2m + k$ försök vilket är betydligt färre än fullständiga 2^k planer. Variabeln k motsvarar hur många så kallade *mittkörningar* som görs. Med mittkörning innebär att alla faktorinställningar körs i origo. Mittkörningarna förbättrar matematiska egenskaperna i en försöksplan eftersom de upprepade försöken låter en testa om modellen passar data.

Då Jones och Nachtsheim introducerade denna gallringsplan år 2011 använde de numerisk sökning av så kallade *D-optimala planer* för att få fram dessa gallringsplaner, se [5]. Sedan dess har Xiao, Lin och Bai introducerat en teoretisk design med vilken man kan komma fram till samma definitiva gallringsplaner som Jones och Nachtsheim, se [12]. Xiao, Lin och Bai använder sig av *konferensmatriser* för att bilda gallringsplanen. Denna strategi har sedan dess blivit standarden då man bildar definitiva gallringsplaner och kommer även användas i denna avhandling.

Definition 4.1. För $m \geq 2$ sägs en $m \times m$ matris \mathbf{C} vara en *konferensmatris* om:

$$\mathbf{C}^T \mathbf{C} = (m - 1) \mathbf{I}_{m \times m}, \quad (4.1)$$

där $\mathbf{C}_{ii} = 0, (i = 1, 2, \dots, m)$ och $\mathbf{C}_{ij} \in \{-1, 1\}, (i \neq j, i, j = 1, 2, \dots, m)$. Konferensmatriser för *jämna* m finns givet i [12, s.1].

Definition 4.2. En *definitiv gallringsplan* är en försöksplan vars designmatris är följande:

$$\mathbf{D} = \begin{pmatrix} \mathbf{C} \\ -\mathbf{C} \\ \mathbf{0} \end{pmatrix}, \quad (4.2)$$

där \mathbf{C} är en $m \times m$ konferensmatris för m faktorer och $\mathbf{0}$ är $k \times m$ nollmatrisen, där k är antalet mittförsök. En definitiv gallringsplan med $n = 2m + k$ körningar betecknas med *DS.n*. För definitionen av definitiva gallringsplaner med hjälp av konferensmatriser, se [12, s.2]. Designen tar då följande form:

Par	Körning (i)	Faktorer				
		$x_{i,1}$	$x_{i,2}$	$x_{i,3}$...	$x_{i,m}$
1	1	0	± 1	± 1	...	± 1
	2	0	∓ 1	∓ 1	...	∓ 1
2	3	± 1	0	± 1	...	± 1
	4	∓ 1	0	∓ 1	...	∓ 1
3	5	± 1	± 1	0	...	± 1
	6	∓ 1	∓ 1	0	...	∓ 1
\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots
m	$2m - 1$	± 1	± 1	± 1	...	0
	$2m$	∓ 1	∓ 1	∓ 1	...	0
Mittkörning	$2m + 1$	0	0	0	...	0

Tabell 4.1: DSD design

Paren består av två körningar som är varandras motsatta tecken eller så är båda körningarna noll. Exempelvis om x_1 har inställningen $+1$ i körning ett så kommer x_1 ha inställningen -1 i körning två, paren som bildas av dessa två körningar kallas för *övervikande par*.

Teorem 4.3. *En definitiv gallringsplan har ingen aliasing mellan huvudeffekterna och*

1. *interceptet,*
2. *rena kvadrater eller*
3. *tvåfaktor interaktionstermerna.*

Vidare har datamatrixen $\mathbf{X} = [\mathbf{1} \quad \mathbf{D} \quad \mathbf{Q}]$, som svarar mot teoretiska modellen

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_m x_m + \beta_{11} x_1^2 + \dots + \beta_{mm} x_m^2 + \epsilon,$$

linjärt oberoende kolonner om $k \geq 1$, där $\mathbf{Q} \in \mathbb{R}^{m \times m}$ ges av $Q_{ii} = 0$ och $Q_{ij} = 1$, $i \neq j$. Om $k = 1$ mittförsök används, så är \mathbf{X} kvadratisk och inverterbar.

Bevis. Låt \mathbf{X}_0 vara faktorvärden som svarar mot interceptet, \mathbf{X}_1 kolonnen som svarar mot huvudeffekterna, \mathbf{X}_2 kolonnen som svarar mot de rena kvadratiske effekterna och \mathbf{X}_3 kolonnen som svarar mot två faktor interaktionstermerna, det vill säga

$$\begin{aligned} \mathbf{X}_0 &= \begin{bmatrix} \mathbf{1}_{m \times 1} \\ \mathbf{1}_{m \times 1} \\ \mathbf{1}_{k \times 1} \end{bmatrix}, \\ \mathbf{X}_1 &= \begin{bmatrix} \mathbf{C}_{m \times m} \\ -\mathbf{C}_{m \times m} \\ \mathbf{0}_{k \times m} \end{bmatrix}, \\ \mathbf{X}_2 &= \begin{bmatrix} \mathbf{Q}_{m \times m} \\ \mathbf{Q}_{m \times m} \\ \mathbf{0}_{k \times m} \end{bmatrix} \quad \text{och} \\ \mathbf{X}_3 &= \begin{bmatrix} \mathbf{P}_{m \times m(m-1)/2} \\ \mathbf{P}_{m \times m(m-1)/2} \\ \mathbf{0}_{k \times m(m-1)/2} \end{bmatrix}. \end{aligned}$$

Formerna $\mathbf{X}_2 = \begin{bmatrix} \mathbf{Q} \\ \mathbf{Q} \\ \mathbf{0} \end{bmatrix}$ och $\mathbf{X}_3 = \begin{bmatrix} \mathbf{P} \\ \mathbf{P} \\ \mathbf{0} \end{bmatrix}$ följer ur definition 4.1 och (4.2). Anta att konferensmatrisen \mathbf{C} har kolonnerna c och d och att interaktionen ges av $p_i = c_i \cdot d_i$. För $-\mathbf{C}$ är motsvarande kolonner $-c$ och $-d$ och deras interaktion är $-c_i \cdot (-d_i) = c_i \cdot d_i = p_i$. Denna kolonn i \mathbf{X}_3 har därmed strukturen $\begin{bmatrix} p \\ p \\ 0 \end{bmatrix}$ och det samma gäller för varje kolonn i \mathbf{X}_3 .

Påståendena i teorem 4.3 kan då skrivas som:

1. Matrisen \mathbf{X}_1 är ortogonal mot \mathbf{X}_0 .
2. Matrisen \mathbf{X}_1 är ortogonal mot \mathbf{X}_2 .
3. Matrisen \mathbf{X}_1 är ortogonal mot \mathbf{X}_3 .

Nedan visas det första påståendet, de två andra följer analogt:

$$\begin{aligned} \mathbf{X}_0^T \mathbf{X}_1 &= [\mathbf{1}_{m \times 1}^T \quad \mathbf{1}_{m \times 1}^T \quad \mathbf{1}_{1 \times k}^T] \times \begin{bmatrix} \mathbf{C}_{m \times m} \\ -\mathbf{C}_{m \times m} \\ \mathbf{0}_{k \times m} \end{bmatrix} \\ &= \mathbf{1}_{m \times 1}^T \mathbf{C}_{m \times m} - \mathbf{1}_{m \times 1}^T \mathbf{C}_{m \times m} + \mathbf{1}_{1 \times k}^T \times \mathbf{0}_{k \times m} \\ &= \mathbf{0}_{k \times m} \end{aligned}$$

Gällande påståendet att \mathbf{X} är kvadratisk, så är det klart att $\mathbf{X}_4 = [\mathbf{1} \quad \mathbf{D} \quad \mathbf{Q}] = [\mathbf{X}_0 \quad \mathbf{X}_1 \quad \mathbf{X}_2]$ har dimension $(2m + k) \times (2m + 1)$, så \mathbf{X}_4 är kvadratisk om och endast om $k = 1$. Därmed återstår att kontrollera att kolonnerna i \mathbf{X}_4 är linjärt oberoende.

Låt $\mathbf{X}_4 \mathbf{b} = \mathbf{0}$ och multiplicera från vänster med \mathbf{X}_1^T . Då fås enligt påstående 1 och 2 ovan att

$$\begin{aligned} \mathbf{0} &= \mathbf{X}_1^T \mathbf{X}_4 = [\mathbf{0} \quad \mathbf{X}_1^T \mathbf{X}_1 \quad \mathbf{0}] \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} \\ &= 2\mathbf{C}^T \mathbf{C} b_1 = (2m - 2)b_1 = \mathbf{0} \implies b_1 = \mathbf{0}. \end{aligned} \tag{4.3}$$

Vidare fås då

$$\mathbf{0} = \mathbf{X}_0^T \mathbf{X}_4 \mathbf{b} = (2m + k) \cdot b_0 + (2m - 2) \mathbf{1}_{m \times 1}^T \mathbf{b}_2, \quad (4.4)$$

att

$$[\mathbf{X}_0 \quad \mathbf{X}_2] \begin{bmatrix} b_0 \\ \mathbf{b}_2 \end{bmatrix} = \mathbf{0} \iff \mathbf{X}_0 b_0 = - \begin{bmatrix} Q \\ Q \\ 0 \end{bmatrix} \mathbf{b}_2. \quad (4.5)$$

Då $k \geq 1$ fås ur sista raden av (4.5) att $b_0 = 0$ och

$$\begin{aligned} \mathbf{0} &= [1 \quad -1 \quad 0 \quad \dots \quad 0] \mathbf{X}_0 b_0 = - [1 \quad -1 \quad 0 \quad \dots \quad 0] \begin{bmatrix} Q \\ Q \\ 0 \end{bmatrix} \mathbf{b}_2 \\ &= b'_1 - b'_2, \end{aligned}$$

där $\mathbf{b}_2 = \begin{bmatrix} b'_1 \\ b'_2 \\ \vdots \end{bmatrix}$.

På samma vis fås att $b'_3 = b'_2$ etc. så

$$\mathbf{b}_2 = \mathbf{1}_{m \times 1}^T b'_1. \quad (4.6)$$

Då ger (4.4) att $(2m - 2) \cdot m b'_1 = 0$, så ur $m \geq 2$ följer $b'_1 = 0$, det vill säga $\mathbf{b} = \mathbf{0}$, och därmed är \mathbf{X}_4 injektiv. \square

Datamatriisen $\mathbf{X}' = [\mathbf{1} \quad \mathbf{D}]$ har linjärt oberoende kolonner även om $k = 0$. Ur $\mathbf{X}' \begin{bmatrix} b_0 \\ \mathbf{b}_1 \end{bmatrix} = \mathbf{0}$ och (4.3) följer nämligen att $b_0 = 0$ om \mathbf{Q} saknas eller är lika med noll.

Försöksplaner var \mathbf{X} är inverterbar kallas *saturerade*. Det innebär att de skattar lika många parametrar p som det finns försök n , det vill säga $p = n$ i (3.7), så att residualen har $n - p = 0$ frihetsgrader. Vidare är också residualen noll vilket innebär att modellen passar data perfekt:

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} = \mathbf{X}^{-1} \mathbf{y},$$

ger $\mathbf{r} = \mathbf{y} - \mathbf{X} \mathbf{b} = \mathbf{0}$. Kvoten i (3.7) blir därmed $\frac{0}{0}$ och all variansanalys är omöjlig för saturerade försöksplaner. En grov tumregel är att man *överanpassar* modellen

till specifika data om $p > \frac{n}{2}$. Anpassning av en första ordningens modell med en DSD-plan har $p = m + 1$ och $n = 2m + k \approx 2p$, så de är ytterst ekonomiska.

Trots det används DSD-planer var rena kvadratiska effekter är inkluderade rätt ofta i litteraturen. En orsak till detta kan vara att det leder till en EED-plan, som diskuteras i mera detalj i exempel 5.9 nedan.

Kapitel 5

Ortogonal blocking och splitplotplaner

Fastän en fullständigt slumpad försöksplan är fördelaktigt så är det ofta i praktiken svårt eller rentav omöjligt att genomföra. Ofta uppstår detta problem när en variabel är väldigt dyr eller tidskrävande att ändra, exempelvis att ställa in exakt temperatur på en smältugn kan kräva mycket energi och vara tidskrävande.

Då en fullständigt slumpad försöksplan inte är möjlig så behöver vi alternativa planer för att lösa problemet. En vanlig sådan är att införa *blocking* i sin försöksplan. Blocking innebär att försök delas in i block enligt någon gemensam faktor. Sedan körs försöken med den gemensamma faktorn konstant inom blocken.

Blocking brukar delas in i två olika undergrupper, *fixerade blockeffekter* samt *stokastiska blockeffekter*. Fixerade blockeffekter innebär att blockeffekten i modellen kan tolkas som fix, kallas även *deterministiska* blockeffekter. Stokastiska blockeffekter innebär att blockeffekterna kan tolkas som stokastiska. I detta kapitel kommer båda att diskuteras.

5.1 Blockindelade försöksplaner

Då alla faktorer inte nollställs efter varje försök så blir analysen av blockade experiment mer komplicerad än fullständigt randomiserade experiment. I och med att vi inte nollställer alla faktorer vid varje försök så introduceras en *slumpningsrestriktion*

för experimentet. Det innebär att även om vi sedan slumpar vår körordning inom blocken så kommer vi inte att uppnå ett fullständigt randomiserat experiment.

Detta fenomen kallas ofta för *restriktionsfel* (eng. *restriction error*). Då man inför slumpningsrestriktion i sitt experiment är det därmed viktigt att ta i beaktande felet som uppstår vilket görs genom att lägga till en term. Termen som läggs till är den tidigare nämnda blockeffekten, denna brukar betecknas med δ . Hur blockeffekten sedan tolkas beror på vilken typ av tidigare nämnda blockeffekter man jobbar med.

Exempel 5.1. I designen nedan har vi fem variabler som kan anta tre olika värden och tre olika block, de två sista körningarna är mittkörningar.

Block	Körning	x_1	x_2	x_3	x_4	x_5
1	1	0	1	1	1	1
	2	0	-1	-1	-1	-1
2	3	1	0	1	-1	-1
	4	-1	0	-1	1	1
3	5	1	1	0	1	-1
	6	-1	-1	0	-1	1
1	7	1	-1	1	0	1
	8	-1	1	-1	0	-1
2	9	1	-1	-1	1	0
	10	-1	1	1	-1	0
3	11	1	1	-1	-1	1
	12	-1	-1	1	1	-1
1	13	0	0	0	0	0
	14	0	0	0	0	0

Tabell 5.1: Blockindelad design med tre block

Det här är ett exempel på en *definitiv gallringsplan* med blockindelning.

Definition 5.2. Låt n vara antalet försök, p antalet parametrar och B antalet block och låt antal försök i block nummer b vara n_b .

Låt y_i vara observation nummer i , i block $b(i)$, där $1 \leq b(i) \leq B$. Inställningen för faktor x_j vid körning i ges av x_{ij} där $j = 1, \dots, m$ anger blocket.

Första ordningens modell blir då:

$$y_i = \beta_0 + \sum_{j=1}^m \beta_j x_{ij} + \delta_{b(i)} + \epsilon_i, \quad b(i) \in 1, \dots, B, \quad i = 1, \dots, n. \quad (5.1)$$

δ ur denna modell tar i beaktande blockindelningen som vi infört i modellen. Denna kallas för *blockeffekten* och kommer att diskuteras mer i detalj. Feltermen ϵ_i antas vara $\epsilon_i \sim N(\mathbf{0}, \sigma^2)$ samt ϵ_i och ϵ_j antas vara oberoende för $i \neq j$.

Definition 5.3. I matrisform kan (5.1) skrivas som:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\epsilon}. \quad (5.2)$$

Här är \mathbf{y} är $n \times 1$ responsvektorn, \mathbf{X} är $n \times p$ datamatrisen.

I modellen är $\boldsymbol{\delta}$ en $B \times 1$ vektor som representerar blockeffekten, $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_n)^T$ är en $n \times 1$ vektor som representerar feltermen och \mathbf{Z} är en $n \times B$ indikatormatris som anger till vilket block försöket hör.

Indikatormatrisen antar värdet ett i det block som försöket hör till och övriga tar värdet noll. För exempel 5.1 skulle \mathbf{Z} se ut så här:

$$\mathbf{Z} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

5.2 Ortogonala block

I ortogonala block försöksplaner modelleras blockeffekten som *fix*. Detta innebär att δ antas ta ett konstant värde i försöken.

Namnet ortogonala försöksplaner kommer från att man väljer att modellera sin modell så att indikatormatrisen \mathbf{Z} är ortogonal mot de linjära termernas huvudeffekter.

För (5.2) skulle därmed:

$$\mathbf{X}^T \mathbf{Z} = \mathbf{0}.$$

För definitionen av ortogonala block, se [8, 8.4.9].

Exempel 5.4. Det tidigare exemplet 5.1 är *inte* en ortogonal försöksplan av första ordningen, eftersom

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & -1 & -1 & -1 & -1 \\ 1 & 1 & 0 & 1 & -1 & -1 \\ 1 & -1 & 0 & -1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & -1 \\ 1 & -1 & -1 & 0 & -1 & 1 \\ 1 & 1 & -1 & 1 & 0 & 1 \\ 1 & -1 & 1 & -1 & 0 & -1 \\ 1 & 1 & -1 & -1 & 1 & 0 \\ 1 & -1 & 1 & 1 & -1 & 0 \\ 1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{och} \quad \mathbf{Z} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

ger

$$\mathbf{X}^T \mathbf{Z} = \begin{bmatrix} 5 & 5 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \neq \mathbf{0}.$$

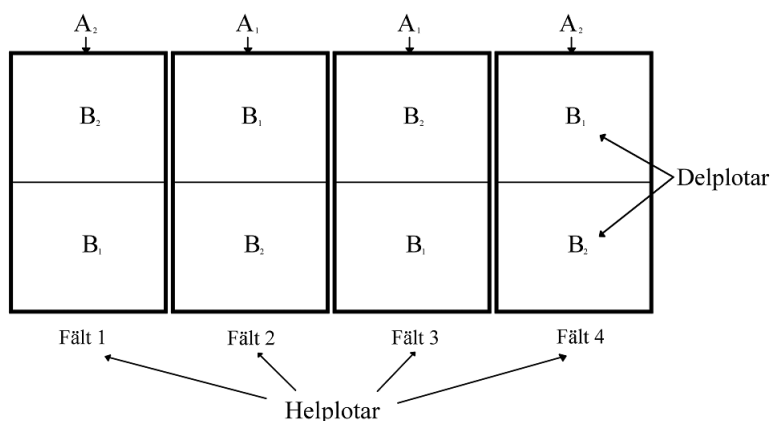
5.3 Splitplotplaner

Splitplotplaner är ett specialfall av blockindelade försöksplaner. Den stora skillnaden från exempelvis ortogonala block försöksplaner är att δ tolkas som *stokastisk*. Detta gör analysen svårare eftersom blockeffekterna i sig ger en ny variansskälla.

Terminologin i splitplotplaner skiljer sig lite från vanlig blockindelning. Blocken kallas för *helplotar* och försöken som körs i blocken delas in i *delplotar*.

Splitplotplaner och dess terminologi har sitt ursprung från jordbruket var bönder ville testa olika typer av bevattningssystem för olika typer av sädeslag. Då bevattningssystemet var rätt svårt att ändra till varje åker så var en fullständigt slumpad försöksplan inte möjlig. Därmed infördes i stället blockindelning på åkrarna, se [8].

Nedan är en simpel splitplot design med fyra helplotar var varje har två delplotar. Variabeln A är svår att ändra medan B är den lätta att ändra.



Figur 5.1: Simpel splitplot design

[4]

Exempel 5.5. Det tidigare exemplet 5.1 kan konstrueras om så att det blir en splitplot design. Anta att x_1 är svår att ändra som vi därmed gärna håller konstant i helplotarna. Denna variabel kommer därmed att slumpas för varje helplot varefter inställningen för de övriga variablerna slumpas in i varje delplot.

HP	SP	x_1	x_2	x_3	x_4	x_5
1	1,1	0	1	1	1	1
	1,2	0	0	1	-1	-1
	1,3	0	-1	1	0	1
	1,4	0	-1	-1	-1	-1
	1,5	0	0	0	0	0
	1,6	0	0	0	0	0
2	2,1	1	1	1	1	1
	2,2	1	0	1	-1	-1
	2,3	1	-1	1	0	1
	2,4	1	-1	-1	-1	-1
3	3,1	-1	1	1	1	1
	3,2	-1	0	1	-1	-1
	3,3	-1	-1	1	0	1
	3,4	-1	-1	-1	-1	-1

Tabell 5.2: Splitplot design med tre helplotar

Då de två mittkörningarna läggs till i helplot ett, leder det till att denna design inte är *balanserad*. Med en *balanserad* design innebär att alla helplotar har lika många delplotar.

5.3.1 Splitplotmodellen

Definition 5.6. Modellen för splitplotplaner är stort sett den samma som definition (5.2). Den viktiga skillnaden är som sagt att $\boldsymbol{\delta}$ tolkas som stokastisk.

Låt n vara antalet försök, p antalet parametrar och B antalet helplotar:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\epsilon}, \quad (5.3)$$

där \mathbf{y} är $n \times 1$ responsvektorn, \mathbf{X} är $n \times p$ modellvektorn, \mathbf{Z} är en $n \times B$ indikatormatris, $\boldsymbol{\delta}$ är $B \times 1$ vektorn som representerar blockeffekten och $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_n)^T$ är $n \times 1$ som representerar delplotfelen.

Här antas att $\boldsymbol{\epsilon} \sim N(\mathbf{0}_n, \sigma^2 \mathbf{I}_n)$ och $\boldsymbol{\delta} \sim N(\mathbf{0}_a, \sigma_h^2 \mathbf{I}_a)$, var \mathbf{I}_n är $n \times n$ identitetsmatrisen. Dessutom antas att $\boldsymbol{\delta}$ och $\boldsymbol{\epsilon}$ är sinsemellan oberoende. För definitionen av splitplotmodellen, se [4].

Exempel 5.7. Låt modellen vara en första ordnings modell som i (5.1). Då blir \mathbf{X} och \mathbf{Z} ur (5.3), för exempel 5.1,

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & -1 & -1 \\ 1 & 0 & -1 & 1 & 0 & 1 \\ 1 & 0 & -1 & -1 & -1 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 0 & 1 \\ 1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix} \quad \text{och} \quad \mathbf{Z} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

5.4 Egenskaper hos en splitplotplan

5.4.1 Kostnad för utförande

Då vi pratar om kostnad gällande försöksplaner så behöver det inte endast handla om pengar utan även tid. I fullständigt randomiserade försöksplaner (eng. *completely randomized design (CRD)*) körs alla tänkbara kombinationer av faktorerna, dessa måste återställas på ett oberoende sätt efter varje körning. Detta leder till noggranna parameterskattningar men är ofta för arbetsdryga för att kunna appliceras i verkligheten.

Anta att vi har en liknande uppsättning som förr med en variabel som är lätt att ändra och en som är svår att ändra. Låt K_h vara kostnaden för att ställa in den variabel som är svår att ändra och K_d vara kostnaden för att ställa in den variabel som är lätt att ändra. Låt vidare B beteckna antalet helplotar och D antalet sammanlagda delplotar.

Kostnaden för en fullständigt randomiserad försöksplan kan då uttryckas som:

$$K_{FRF} = D(K_h + K_d).$$

Eftersom vi i splitplotplaner inte kör alla möjliga kombinationer utan delar upp försöken så som tidigare beskrivet, kan vi på samma vis skriva kostnaden för dem som:

$$K_{SPP} = BK_h + DK_d.$$

Den extra kostnaden som medföljer från en fullständigt randomiserad försöksplan i jämförelse med en splitplotplan blir därmed:

$$\begin{aligned} K_{FRF} - K_{SPP} &= D(K_h + K_d) - (BK_h + DK_d) \\ &= (D - B)K_h. \end{aligned}$$

5.5 Analys av splitplotplaner

5.5.1 Responsens kovarians

Teorem 5.8. Låt modellen vara som i (5.3), kovariansen för \mathbf{y} är då $\Sigma := \sigma_h^2 \mathbf{Z}\mathbf{Z}^T + \sigma^2 \mathbf{I}_n$ och Σ är en inverterbar matris.

Bevis. Låt vår modell vara (5.3):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\epsilon}$$

Variansen för \mathbf{y} kan då beräknas som:

$$\begin{aligned} \text{Var}(\mathbf{y}) &= \text{Var}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\epsilon}) \\ &= \text{Var}(\mathbf{X}\boldsymbol{\beta}) + \text{Var}(\mathbf{Z}\boldsymbol{\delta}) + \text{Var}(\boldsymbol{\epsilon}) + 2\text{Cov}(\mathbf{X}\boldsymbol{\beta}, \mathbf{Z}\boldsymbol{\delta}) + 2\text{Cov}(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\epsilon}) + 2\text{Cov}(\mathbf{Z}\boldsymbol{\delta}, \boldsymbol{\epsilon}) \\ &= \text{Var}(\mathbf{Z}\boldsymbol{\delta}) + \text{Var}(\boldsymbol{\epsilon}) \\ &= \mathbf{Z}\text{Var}(\boldsymbol{\delta})\mathbf{Z}^T + \sigma^2 \mathbf{I}_n \\ &= \mathbf{Z}\sigma_h^2 \mathbf{I}_a \mathbf{Z}^T + \sigma^2 \mathbf{I}_n \\ &= \sigma_h^2 \mathbf{Z}\mathbf{Z}^T + \sigma^2 \mathbf{I}_n. \end{aligned}$$

Eftersom Σ består av en summa mellan en positivt definit matris och en positivt semidefinit matris så är Σ själv en positivt definit matris och därmed inverterbar.

□

I teorem 5.8 bevisades att Σ för splitplot ges av $\sigma_h^2 \mathbf{Z}\mathbf{Z}^T + \sigma^2 \mathbf{I}_n$, därmed är det nödvändigt att använda generaliserade minsta kvadratmetoden som beskrivet i kapitel 3.2.

Problemet som uppstår är att Σ är okänd eftersom σ och σ_h i teorem 5.8 inte är kända. För att lösa detta substitueras estimaten $\hat{\sigma}^2$ och $\hat{\sigma}_h^2$ av σ^2 och σ_h in i (3.9) vilket på samma vis som i definition 3.2 ger den så kallade *möjliga generaliserade minstakvadratskattningen* av parametervektorn β :

$$\mathbf{b}_{FGLS} = (\mathbf{X}^T \widehat{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \widehat{\Sigma}^{-1} \mathbf{y},$$

där

$$\widehat{\Sigma} = \hat{\sigma}_h^2 \mathbf{Z}\mathbf{Z}^T + \hat{\sigma}^2 \mathbf{I}_n.$$

För att få skattningarna $\hat{\sigma}_h^2$ och $\hat{\sigma}^2$ rekommenderar Jones & Nachtshiem [4] *residual maximum likelihood (REML)*. Denna rekommendation baserar sig på analys som Letsinger et al. gjorde på bland annat splitplotplaner i [6].

5.5.2 Maximum likelihood för splitplotplaner

Om vi låter modellen för splitplot vara given som i (5.3) så kan ℓ funktionen som definierad i (3.2) användas:

$$\ell = \log(L) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma| - \frac{1}{2} (\mathbf{y} - \mathbf{X}\beta)^T \Sigma^{-1} (\mathbf{y} - \mathbf{X}\beta), \quad (5.4)$$

där $\Sigma = \sigma_h^2 \mathbf{Z}\mathbf{Z}^T + \sigma^2 \mathbf{I}_n$, se teorem 5.8.

Då formen för Σ i splitplot är känd kan maximum likelihood lösningarna härledas längre än vi tidigare visat i kapitel 3. Målet är därmed att maximera funktionen (5.4) vilket görs genom att söka nollställen till partiella derivatorna med avseende på σ_h^2 och σ^2 .

Maximum likelihood-skattningen för σ_h^2 och σ^2 i splitplot.

För att beräkna partiella derivatan av ℓ med avseende på σ_h^2 och σ^2 behövs två följande räkneregler för partiell derivata, se [10, Appendix M.7].

$$\begin{aligned} \frac{\partial \mathbf{A}(x)^{-1}}{\partial x} &= -\mathbf{A}(x)^{-1} \frac{\partial \mathbf{A}(x)}{\partial x} \mathbf{A}(x)^{-1} \\ \frac{\partial}{\partial x} \log |\mathbf{A}(x)| &= \text{tr} \left(\mathbf{A}(x)^{-1} \frac{\partial \mathbf{A}(x)}{\partial x} \right), \end{aligned}$$

där $tr(\mathbf{A})$ står för spåret av matrisen \mathbf{A} , se [10, Appendix M.6]. Partiella derivatan av ℓ med avseende på σ_h^2 ges av:

$$\ell_{\sigma_h^2} = \frac{\partial \ell}{\partial \sigma_h^2} = -\frac{1}{2} tr \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \sigma_h^2} \right) + \frac{1}{2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \sigma_h^2} \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}).$$

Givet att $\frac{\partial \boldsymbol{\Sigma}}{\partial \sigma_h^2} = \mathbf{Z}\mathbf{Z}^T$ fås

$$\ell_{\sigma_h^2} = -\frac{1}{2} tr(\boldsymbol{\Sigma}^{-1} \mathbf{Z}\mathbf{Z}^T) + \frac{1}{2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T \boldsymbol{\Sigma}^{-1} \mathbf{Z}\mathbf{Z}^T \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}). \quad (5.5)$$

För $\ell_{\sigma^2} = \frac{\partial \ell}{\partial \sigma^2}$ gäller, då $\frac{\partial \boldsymbol{\Sigma}}{\partial \sigma^2} = \mathbf{I}_n$, att

$$\ell_{\sigma^2} = -\frac{1}{2} tr(\boldsymbol{\Sigma}^{-1}) + \frac{1}{2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T \boldsymbol{\Sigma}^{-2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}). \quad (5.6)$$

Maximum likelihood lösningarna till σ_h och σ ges av nollställena till (5.5) och (5.6). Då estimatet för σ_h^2 substitueras in fås ML lösningarna för σ_h^2

$$tr(\widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{Z}\mathbf{Z}^T) = (\mathbf{y} - \mathbf{X}\mathbf{b})^T \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{Z}\mathbf{Z}^T \widehat{\boldsymbol{\Sigma}}^{-1} (\mathbf{y} - \mathbf{X}\mathbf{b}), \quad (5.7)$$

och på samma vis för σ^2 :

$$tr(\widehat{\boldsymbol{\Sigma}}^{-1}) = (\mathbf{y} - \mathbf{X}\mathbf{b})^T \widehat{\boldsymbol{\Sigma}}^{-2} (\mathbf{y} - \mathbf{X}\mathbf{b}). \quad (5.8)$$

Dessa två uttryck kan göras lite algebraiskt lättare. Givet (3.4) fås att

$$\begin{aligned} \mathbf{X}^T \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{X}\mathbf{b} &= \mathbf{X}^T \widehat{\boldsymbol{\Sigma}} \mathbf{y} \\ \implies \mathbf{b} &= (\mathbf{X}^T \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{y} \\ \implies \widehat{\boldsymbol{\Sigma}}^{-1} (\mathbf{y} - \mathbf{X}\mathbf{b}) &= (\widehat{\boldsymbol{\Sigma}}^{-1} - \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{X} (\mathbf{X}^T \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \widehat{\boldsymbol{\Sigma}}^{-1}) \mathbf{y}. \end{aligned}$$

Definiera sedan följande matris

$$\mathbf{P} = \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1} \mathbf{X} (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\Sigma}^{-1}. \quad (5.9)$$

vilket då vi substituerar in $\widehat{\mathbf{P}}$ för \mathbf{P} slutligen ger

$$\widehat{\boldsymbol{\Sigma}}^{-1} (\mathbf{y} - \mathbf{X}\mathbf{b}) = \widehat{\mathbf{P}} \mathbf{y}.$$

Då kan (5.7) och (5.8) förenklas till

$$tr(\widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{Z}\mathbf{Z}^T) = \mathbf{y}^T \widehat{\mathbf{P}} \mathbf{Z}\mathbf{Z}^T \widehat{\mathbf{P}} \mathbf{y} \quad \text{och} \quad (5.10)$$

$$tr(\widehat{\boldsymbol{\Sigma}}^{-1}) = \mathbf{y}^T \widehat{\mathbf{P}}^2 \mathbf{y}. \quad (5.11)$$

5.5.3 REML för splitplotplaner

För att använda REML måste responsvektorn \mathbf{y} transformeras som beskrivet i avsnitt 3.3. Låt modellen vara som (5.3):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\epsilon}.$$

Välj en vektor \mathbf{r} sådan att i

$$\mathbf{r}^T \mathbf{y} = \mathbf{r}^T \mathbf{X}\boldsymbol{\beta} + \mathbf{r}^T \mathbf{Z}\boldsymbol{\delta}$$

är

$$\mathbf{r}^T \mathbf{X}\boldsymbol{\beta} = \mathbf{0} \quad \text{för alla } \boldsymbol{\beta}. \quad (5.12)$$

Då kan REML-ekvationerna härledas ur (5.10) och (5.11) genom att substituera in matrisen $\widetilde{\mathbf{M}}$ som är beskriven i (3.13). Detta görs genom följande substitutioner

$$\begin{aligned} \mathbf{y} &\rightarrow \widetilde{\mathbf{M}}\mathbf{y} \\ \mathbf{X} &\rightarrow \widetilde{\mathbf{M}}\mathbf{X} = \mathbf{0} \\ \mathbf{Z} &\rightarrow \widetilde{\mathbf{M}}\mathbf{Z} \\ \widehat{\boldsymbol{\Sigma}} &\rightarrow \widetilde{\mathbf{M}}\widehat{\boldsymbol{\Sigma}}\widetilde{\mathbf{M}}^T. \end{aligned}$$

Dessa substitutioner leder till att matrisen $\widehat{\mathbf{P}}$ i (5.9) med (5.12) tar formen

$$\widehat{\mathbf{P}} = (\widetilde{\mathbf{M}}\widehat{\boldsymbol{\Sigma}}\widetilde{\mathbf{M}}^T)^{-1},$$

där

$$\widehat{\boldsymbol{\Sigma}} = \hat{\sigma}_h^2 \mathbf{Z}\mathbf{Z}^T + \hat{\sigma}^2 \mathbf{I}_n. \quad (5.13)$$

Detta ger de två motsvarande REML lösningarna för σ_h och σ till (5.10) och (5.11)

$$\begin{aligned} tr(\widehat{\mathbf{P}}\widetilde{\mathbf{M}}\mathbf{Z}(\widetilde{\mathbf{M}}\mathbf{Z})^T) &= (\widetilde{\mathbf{M}}\mathbf{y})^T \widehat{\mathbf{P}}\widetilde{\mathbf{M}}\mathbf{Z}(\widetilde{\mathbf{M}}\mathbf{Z})^T \widehat{\mathbf{P}}\widetilde{\mathbf{M}}\mathbf{y} \\ \iff tr(\widehat{\mathbf{P}}\widetilde{\mathbf{M}}\mathbf{Z}\mathbf{Z}^T\widetilde{\mathbf{M}}^T) &= \mathbf{y}^T \widetilde{\mathbf{M}}^T \widehat{\mathbf{P}}\widetilde{\mathbf{M}}\mathbf{Z}\mathbf{Z}^T\widetilde{\mathbf{M}}^T \widehat{\mathbf{P}}\widetilde{\mathbf{M}}\mathbf{y} \end{aligned} \quad (5.14)$$

$$\begin{aligned} tr(\widehat{\mathbf{P}}) &= (\widetilde{\mathbf{M}}\mathbf{y})^T \widehat{\mathbf{P}}^2 \widetilde{\mathbf{M}}\mathbf{y} \\ &= \mathbf{y}^T \widetilde{\mathbf{M}}^T \widehat{\mathbf{P}}^2 \widetilde{\mathbf{M}}\mathbf{y}. \end{aligned} \quad (5.15)$$

I praktiken så hittas lösningarna till (5.14) och (5.15) genom att lösa det olinjära algebraiska ekvationssystemet i σ_h^2 och σ^2 med hjälp av numeriska metoder, se [10, Kap.8].

5.6 Splitplot och DSD

En metod för att använda definitiva gallringsplaner inom en splitplot struktur är den så kallade *minimala helplotgrupperingsmetoden* (eng. *minimum whole-plot grouping method*). Detta innebär att för $B \geq 2$, var B står för antalet helplotar, så kan körningarna sorteras om så att antalet helplotar blir så få som möjligt. Detta är fördelaktigt eftersom det minimerar kostnaden relaterad till att byta inställningen för helplotfaktorer, vilket är den bakomliggande idén till splitplot.

Följande är ett exempel var splitplotplaner kombineras med DSD. Responsdata och designmatrisen är taget ur [7], men jag kommer presentera både fallet då vi har en EED-plan och då vi inte har det.

Exempel 5.9. Nedan är en *DS.17* design med två helplotsfaktorer och sex delplotsfaktorer, var H_1 och H_2 är helplotsfaktorerna och $S_1, S_2, S_3, S_4, S_5, S_6$ är delplotsfaktorerna.

HP	H_1	H_2	S_1	S_2	S_3	S_4	S_5	S_6
1	-1	-1	1	-1	1	0	-1	1
	-1	-1	1	1	-1	1	0	-1
	-1	-1	-1	1	1	-1	1	0
2	-1	0	-1	-1	-1	1	1	1
3	-1	1	0	1	-1	-1	-1	1
	-1	1	-1	0	1	1	-1	-1
	-1	1	1	-1	0	-1	1	-1
4	0	1	1	1	1	1	1	1
5	0	0	0	0	0	0	0	0
6	0	-1	-1	-1	-1	-1	-1	-1
7	1	-1	0	-1	1	1	1	-1
	1	-1	1	0	-1	-1	1	1
	1	-1	-1	1	0	1	-1	1
8	1	0	1	1	1	-1	-1	-1
9	1	1	-1	1	-1	0	1	-1
	1	1	-1	-1	1	-1	0	1
	1	1	1	-1	-1	1	-1	0

Tabell 5.3: DS.17 design för splitplot

Låt modellen vara som i (5.3), det vill säga

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\epsilon}.$$

Låt oss sedan anpassa en fullständig första ordningens modell som även innehåller rena kvadratiska effekter,

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_m x_m + \beta_{11} x_1^2 + \dots + \beta_{mm} x_m^2 + \epsilon.$$

Datamatriken definieras som $\mathbf{X} = [\mathbf{1} \ \mathbf{D} \ \mathbf{Q}]$, där \mathbf{D} är designmatriken för definitiva gallringsplaner som beskrivet i definition 4.2 och \mathbf{Q} är de rena kvadratiska effekterna, som beskrivet i teorem 4.3. Orsaken till valet av denna modell är att det leder till en saturerad försöksplan vilket enligt korollarium 3.9 leder till en EED-plan.

Matrisen \mathbf{Z} , som beskriven i definition 5.6, tar formen

$$\mathbf{Z}^T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

Vidare låt responsvektorn \mathbf{y} ur Lin och Yang [7] vara given av

$$\mathbf{y}^T = [55,1 \ 56,4 \ 50,5 \ 50,3 \ 58,0 \ 49,6 \ 55,0 \ 48,8 \ \dots \\ 49,0 \ 42,7 \ 50,7 \ 51,8 \ 41,4 \ 48,2 \ 41,7 \ 42,9 \ 52,1] .$$

Eftersom det är en EED-plan så kan ordinära minstakvadratmetoden användas som beskrivet i avsnitt 2.1 för att få estimat av helplots- och delplotseffekterna.

Effekt	Estimat	Effekt	Estimat
Intercept	48,955		
\hat{H}_1	-3,305	\hat{H}_1^2	4,052
\hat{H}_2	-0,019	\hat{H}_2^2	0,525
\hat{S}_1	3,437	\hat{S}_1^2	-4,525
\hat{S}_2	-0,275	\hat{S}_2^2	-0,876
\hat{S}_3	-0,508	\hat{S}_3^2	1,584
\hat{S}_4	0,004	\hat{S}_4^2	1,435
\hat{S}_5	0,120	\hat{S}_5^2	0,158
\hat{S}_6	0,290	\hat{S}_6^2	-1,500

Tabell 5.4: OLS-estimat för helplotseffekterna och delplotseffekterna

Ifall \mathbf{Q} stryks ur modellen fås icke längre en EED-plan. Det här gör analysen avsevärt mera omständig. Då bör ordinära minstakvadratmetoden inte längre användas, utan generaliserade minstakvadratmetoden tillsammans med REML som beskrivet i avsnitt 5.5.3 bör användas.

Med hjälp av en numerisk lösare i MATLAB kan $\hat{\sigma}_h^2$ och $\hat{\sigma}^2$ lösas ur (5.14) och (5.15), för vilket koden hittas i bilaga. Vi får lösningen

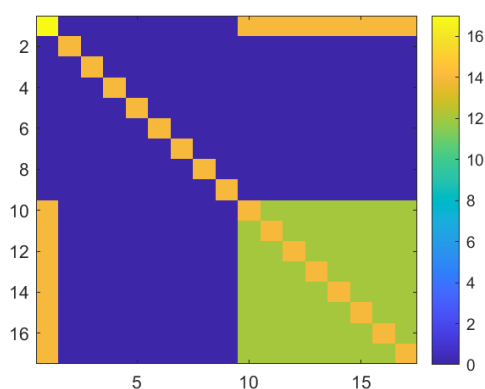
$$\begin{cases} \hat{\sigma}_h^2 = 4,4825 \\ \hat{\sigma}^2 = 4,1295. \end{cases}$$

Till följande används dessa estimat i definition 3.2, med Σ givet av (5.13) för att få följande estimat \mathbf{b}_{GLS} :

Effekt	Estimat
Intercept	49,225
\hat{H}_1	-3,305
\hat{H}_2	-0,019
\hat{S}_1	3,437
\hat{S}_2	-0,275
\hat{S}_3	-0,508
\hat{S}_4	0,004
\hat{S}_5	0,120
\hat{S}_6	0,290

Tabell 5.5: GLS-estimat för helplotseffekterna och delplotseffekterna

Vi ser att bortsett från interceptet så är estimaten de samma som i ordinära fallet. Orsaken till att interceptet inte är samma ses genom att bilda informationsmatrisen $\mathbf{X}^T \mathbf{X}$.



Figur 5.2: Informationsmatrisen $\mathbf{X}^T \mathbf{X}$

I figur 5.2 ses att det finns aliasing mellan interceptet och de kvadratiska effekterna vilket förklarar att interceptet inte blir samma för \mathbf{b} och \mathbf{b}_{GLS} . Avslutningsvis ses att för de kvadratiska effekterna, det vill säga 10 – 17 i bilden, finns det stark aliasing mellan estimaten. Därmed bör kvadratiska effekterna inte användas som estimat även om de var inkluderade i modellen för att bilda en EED-försöksplan.

Detta exempel inspirerade korollarium 3.9 och insikten om att DSD-planen $\mathbf{X} = [\mathbf{1} \ \mathbf{D} \ \mathbf{Q}]$ är saturerad då $k = 1$.

Litteraturförteckning

- [1] Jacob Albrecht. <https://www.mathworks.com/matlabcentral/fileexchange/50057-definitive-screening-designs>, MATLAB Central File Exchange., Definitive Screening Designs, 2022.
- [2] Virgil L Anderson and Robert A McLean. *Design of experiments: A Realistic Approach*. CRC Press, 2018.
- [3] Julian James Faraway. *Practical regression and ANOVA using R.*, volume 168. University of Bath Bath, 2002.
- [4] Bradley Jones and Christopher Nachtsheim. Split-plot designs: What, why, and how. *Journal of Quality Technology*, 41, 10 2009.
- [5] Bradley Jones and Christopher J. Nachtsheim. A class of three-level designs for definitive screening in the presence of second-order effects. *Journal of Quality Technology*, 43(1):1–15, 2011.
- [6] Jennifer D. Letsinger, Raymond H. Myers, and Marvin Lentner. Response surface methods for bi-randomization structures. *Journal of Quality Technology*, 28(4):381–397, 1996.
- [7] Chang-Yun Lin and Po Yang. Response surface methodology using split-plot definitive screening designs. *Journal of Quality Technology*, 47(4):351–362, 2015.
- [8] Raymond H Myers, Douglas C Montgomery, and Christine M Anderson-Cook. *Response surface methodology: process and product optimization using designed experiments*. John Wiley & Sons, 2016.
- [9] G.G. Roussas. *A Course in Mathematical Statistics*. Elsevier Science, 1997.
- [10] Shayle R Searle, George Casella, and Charles E McCulloch. *Variance components*. John Wiley & Sons, 2009.

- [11] G. Geoffrey Vining, Scott M. Kowalski, and Douglas C. Montgomery. Response surface designs within a split-plot structure. *Journal of Quality Technology*, 37(2):115–129, 2005.
- [12] Lili Xiao, Dennis K. J. Lin, and Fengshan Bai. Constructing definitive screening designs using conference matrices. *Journal of Quality Technology*, 44(1):2–8, 2012.

Bilaga A

MATLAB-kod för exempel 5.9.

Nedan hittas MATLAB-koden för exempel 5.9. Koden använder [1] för att bilda DSD-planer varefter resten är min egen kod som får användas helt fritt, under licensen public domain. Först byggs modellen upp så att den är samma som i [7].

```
%Här bygger vi upp modellen så att den är samma
%som i Lin&Yang
clear variables

m=8 % Antalet kvantitativa variabler
HP=2 % Antalet helplotar

% DSD-plan, för koden till denna se litteraturförteckning
D=dsd(m,0,1);

D=sortrows(D); %sorterar raderna som splitplot
N=size(D,1) %DS.N

%Ändrar om så att D ser likadan ut som i Lin&Yang
D=D([1 2 3 4 5 6 7 10 9 8 13 12 11 14 17 16 15],:);
D=D(:, [1 2 5 7 8 4 3 6]);

%matar in y-värdena
y=[55.073;56.359;50.529;50.349;58.019;49.619;55.049;48.806;
  48.955;42.708;50.650;51.752;41.401;48.219;41.676;42.943;
  52.089];
```

Följande kod bygger upp matrisen Z och skulle enkelt kunna formuleras om som en generisk funktion, till exempel $Z=\text{byggZ}(D,HP)$.

```
%bygger upp hur många helplotar det finns ,
%detta för HP helplotsfaktorer
wp=D(:,1:HP)
wpDiff=diff(wp) %räknar skillnaden mellan raderna på wp

%lagar ny kolumn med nollor
wpNewCol=zeros(length(wp),1);
%sätter första raden i kolumnen till ett
wpNewCol(1) = 1 ;

%bygger upp vilken WP en körning hör till
for ii = 2:length(wpNewCol)
    if ~ any(wpDiff(ii-1, :))
        wpNewCol(ii) = wpNewCol(ii - 1) ;
    else
        wpNewCol(ii) = wpNewCol(ii -1) + 1;
    end
end

new_wp = [wp, wpNewCol] %sista columnen anger nu vilken HP

%Bygger upp Z
m=max(wpNewCol) % Antalet WP
n=length(wpNewCol)/m % Antalet SP per WP
Z=zeros(m*n,m);
for k=1:m*n
    Z(k,wpNewCol(k))=1;
end
Z
```

Till följande kontrolleras att modellen är en EED-plan då vi låter $\mathbf{X} = [\mathbf{1} \quad \mathbf{D} \quad \mathbf{Q}]$ varefter ordinära minstakvadratmetoden används för att få estimat av β .

%OLS estimaten för exemplet.

```
%första ordningens modell+kvadratiska effekter
X2=[ones(N,1) D D.^2]
```

%Testar att det är en EED-plan så vi kan använda OLS

```
K=(X2' * X2)^(-1) * X2' * (Z * Z') * X2
A=X2 * K
B=Z * Z' * X2
```

```
%Kör OLS
regress(y, X2)
```

Slutligen anpassas i stället modellen var vi exkluderar de kvadratiska effekterna till data, vilket innebär att vi inte längre har en EED-plan och REML bör användas tillsammans med generaliserade minstakvadratmetoden för att få estimat av β .

```
% REML
```

```
%x(1)=sigma_h
%x(2)=sigma
```

```
X=[ones(N,1) D] %första ordningens modell
P=X*(X'*X)^(-1)*X'
n=size(P,1)
p=size(X,2)-1
M= eye(n)-P
```

```
%Bör nu plocka ut n-(p+1) rader ur M
M_tilde=M(1:n-p-1,:)
```

```
%kovariansmatrisen
Sigma_hatt=@(x) x(1)*Z*Z'+x(2)*eye(n);
%matrisen P
P_hatt=@(x) (M_tilde*Sigma_hatt(x)*M_tilde')^(-1)
```

```
%Bildar första funktionen i ekvationssystemet
A1=@(x) trace(P_hatt(x)*M_tilde*Z*Z'*M_tilde')
A2=@(x) y'*M_tilde'*P_hatt(x)*M_tilde*Z*Z'...*
M_tilde'*P_hatt(x)*M_tilde*y
A3=@(x) A1(x)-A2(x)
```

```
%Bildar andra funktionen i ekvationssystemet
B1=@(x) trace(P_hatt(x))
B2=@(x) y'*M_tilde'*P_hatt(x)^2*M_tilde*y
B3=@(x) B1(x)-B2(x)
```

```
f=@(x) [A3(x);B3(x)]
varianser=fsolve(f,[1,1])
var_h=varianser(1);
var_s=varianser(2);

f(varianser)    % Kollar att det är en äkta lösning

%bGLS estimaten bGLS=(X'*Sigma_hatt(varianser)^-1*X)...
^-1*X'*Sigma_hatt(varianser)^-1*y

%Ritar ut informationsmatrisen
info=X'*X
imagesc(info)
colorbar
C=inv(info)
```

Även denna REML-kod kan med små ändringar göras till en fristående generisk funktion `[bGLS,var_h,var_s]=REML(X,Z,y)`.