**Simulated annealing write-up**

This write up is on using the simulated annealing algorithm to find the optimal designs.

Simulated annealing is a good method to solve the global optimization problem of locating a good approximation of the global optimum of a given function in a large space.

During the process of optimization of this method, the worsen case compare to the previous case may be accepted based on the acceptance probability defined by the user.

Despite using the optim function, the optimization process is still very sensitive based on the temperature given by the user. This sensitivity issue can be noted when comparing the average efficiency factors, because the differences of the average efficiency factors can be very small.

Hence, I believe I still need to initialise the temperature using the method described in Jansen et al. (1992) for the simulated annealing algorithm. Their method first perform a simulated annealing algorithm of a given design with a very high temperature of 10e9, no cooling-down, acceptance probability of 0.5 and 20000 iterations. During this optimization process, if a worse case compare to the previous case has occurred and this worse case is still been accepted based on the temperature and the acceptance probability, we will record the differences in their value generated from the objective function. After this optimization process is finished, a better temperature can be obtained from the mean of differences and divided by the acceptance probability of 0.5, in this case.

The current simulated annealing algorithm search from the optimal design by random swapping any two rows of the treatment design matrix and maximised the average efficiency factors.

During the assignment for our MudPIT-iTRAQ experiment, note that number of the animals from Phase 1 need to be greater or equal to the number of tags. This is to avoid all the degrees of freedom from the animals to be confounded with the tags.

**Example 1: assigning completely randomised design to randomised bock design**

**(Phase 1: 4 animals, 2 treatments, Phase 2: 4 MS runs and 4-plex iTRAQ tag system)**

Phase 1 design consists of 4 animals randomly assigned to 2 treatment groups: controlled and diseased groups, where animals A and C are assigned to controlled group and animals B and D are assigned to diseased group.

The theoretical ANOVA table showing the coefficients of the variance components for the Phase 1 experiment can be written as

DF Ani

Between Ani

Trt 1 1

Residual 2 1

The design for the second phase experiment aims to assign the animal in the first phase to the block structure in the second phase.

Suppose the second phase experiment uses 4 MS runs and 4-plex iTRAQ tag system.

The aim of the animal allocation is to minimise the level of confounding between animal effects with both runs and tag effects.

Another way of describing the above point is to maximise the amount of information in the within runs and tags stratum, hence the second phase design can be considered as a row-column design.

The orthogonal projector for the within runs and tags stratum can be written as

The information matrix can be written as

where denotes the animal design matrix.

The criterions are to find that maximises the average efficiency factor and the *numbers* of the canonical efficiency factors.

describe how I use the optim function of R….

discuss the paramters…

Based on the current example and the criteria described, the best animal design was found to be the one below

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Run | Tag | | | |
| 1 | 2 | 3 | 4 |
| 1 | D | C | B | A |
| 2 | A | B | C | D |
| 3 | B | D | A | C |
| 4 | C | A | D | B |

The treatment design for Phase 2 experiment can be made from the animal and treatment allocation of the Phase 1 experiment as shown below

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Run | Tag | | | |
| 1 | 2 | 3 | 4 |
| 1 | Diseased | Controlled | Diseased | Controlled |
| 2 | Controlled | Diseased | Controlled | Diseased |
| 3 | Diseased | Diseased | Controlled | Controlled |
| 4 | Controlled | Controlled | Diseased | Diseased |

The theoretical ANOVA table for this two-phase experiment with the animal and treatment allocation described can be shown as below

$ANOVA

DF e Ani Run

Between Run 3 1 0 4

Within

Between Ani

Trt 1 1 4 0

Residual 2 1 4 0

Residual

Tag 3 1 0 0

Residual 6 1 0 0

$EF

Trt Tag eff.Trt eff.Tag

Between Run

Within

Between Ani

Trt 8 1

Residual

Tag 4 1

This theoretical ANOVA table shows all the animal and treatment information are in the within runs stratum from the optimisation procedure described here. In addition, tag does not confound with the animals.

**Example 2: assigning completely randomised design to randomised bock design**

**(Phase 1: 6 animals, 2 treatments, Phase 2: 6 MS runs and 4-plex iTRAQ tag system)**

For this case, the optimisation process obtained the best average efficiency factor for the animal information in the within runs stratum is 0.8936755. This average efficiency factor is calculated from the harmonic mean of the canonical efficiency factors of 1.0000, 0.9375, 0.9375, 0.8125 and 0.8125. This result is obtained from the temp of 0.06842407 with 5000 iterations.

The eigenvectors, which corresponds to the canonical efficiency factors described above, are

[,1] [,2] [,3] [,4] [,5]

[1,] -0.4082483 0.4337891 -0.3809991 -0.50534565 -0.2792116

[2,] 0.4082483 0.4337891 -0.3809991 0.50534565 0.2792116

[3,] 0.4082483 0.1130603 0.5661720 -0.01086849 -0.5772480

[4,] 0.4082483 -0.5468495 -0.1851729 -0.49447715 0.2980364

[5,] -0.4082483 -0.5468495 -0.1851729 0.49447715 -0.2980364

[6,] -0.4082483 0.1130603 0.5661720 0.01086849 0.5772480

Based on the current example and the criteria described, the best animal design was found to be the one in a matrix notation below. In this matrix, the rows denote the MS runs and columns denotes the tags.

[,1] [,2] [,3] [,4]

[1,] "D" "E" "F" "C"

[2,] "A" "B" "D" "F"

[3,] "F" "A" "B" "E"

[4,] "B" "C" "A" "D"

[5,] "C" "D" "E" "B"

[6,] "E" "F" "C" "A"

> with(design.df, table(Ani, Run))

Run

Ani 1 2 3 4 5 6

A 1 1 0 1 0 1

B 1 0 1 1 0 1

C 1 0 1 1 1 0

D 0 1 1 0 1 1

E 0 1 0 1 1 1

F 1 1 1 0 1 0

> N = with(design.df, table(Ani, Run))

> N %\*% t(N)

Ani

Ani A B C D E F

A 4 3 2 2 3 2

B 3 4 3 2 2 2

C 2 3 4 2 2 3

D 2 2 2 4 3 3

E 3 2 2 3 4 2

F 2 2 3 3 2 4

The treatment design for Phase 2 experiment can be made from the animal and treatment allocation of the Phase 1 experiment as shown below

[,1] [,2] [,3] [,4]

[1,] "controlled" "controlled" "controlled" "diseased"

[2,] "controlled" "diseased" "controlled" "diseased"

[3,] "controlled" "controlled" "diseased" "diseased"

[4,] "diseased" "diseased" "diseased" "controlled"

[5,] "diseased" "controlled" "diseased" "controlled"

[6,] "diseased" "diseased" "controlled" "controlled"

The theoretical ANOVA table for this two-phase experiment with the animal and treatment allocation described can be shown as below

$ANOVA

DF e Ani Run

Between Run

Between Ani

Trt 1 1 1/2 4

Residual 3 1 1/2 4

Residual 1 1 0 4

Within

Between Ani

Trt 1 1 81/22 0

Residual 4 1 315/88 0

Residual

Tag 3 1 0 0

Residual 10 1 0 0

$EF

Trt Tag eff.Trt eff.Tag

Between Run

Between Ani

Trt 1 1/12

Residual

Within

Between Ani

Trt 11 11/12

Residual

Tag 6 1

Note this optimisation process, 0.8936755 of animal information is found in the within runs stratum, hence it is expected that treatment information may also be separated across different strata. This theoretical ANOVA table shows 11/12 of treatment information is in the within runs stratum. In addition, the coefficients of animal variance components are not identical for the treatment and the residual in the between animals within runs stratum. Hence, to test for the differences between the treatment groups, the coefficients of animal variance components needs to be adjusted.

**Example 3: assigning completely randomised design to randomised bock design**

**(Phase 1: 9 animals, 3 treatments, Phase 2: 9 MS runs and 4-plex iTRAQ tag system)**

For this case, the optimisation process obtained the best average efficiency factor for the animal information in the within runs stratum is 0.8340019. This average efficiency factor is calculated from the harmonic mean of the canonical efficiency factors of 0.9924616, 0.9924616, 0.853294, 0.853294, 0.7792444, 0.7792444, 0.75 and 0.75. This result is obtained from the temp of 0.02416828 with 10000 iterations

Based on the current example and the criteria described, the best animal design was found to be the one in a matrix notation below. In this matrix, the rows denote the MS runs and columns denotes the tags.

[,1] [,2] [,3] [,4]

[1,] "D" "B" "C" "H"

[2,] "A" "C" "E" "D"

[3,] "I" "A" "G" "B"

[4,] "G" "I" "H" "C"

[5,] "F" "D" "A" "I"

[6,] "C" "G" "F" "E"

[7,] "B" "F" "D" "G"

[8,] "H" "E" "I" "F"

[9,] "E" "H" "B" "A"

> (N = with(design.df, table(Ani, Run)))

Run

Ani 1 2 3 4 5 6 7 8 9

A 0 1 1 0 1 0 0 0 1

B 1 0 1 0 0 0 1 0 1

C 1 1 0 1 0 1 0 0 0

D 1 1 0 0 1 0 1 0 0

E 0 1 0 0 0 1 0 1 1

F 0 0 0 0 1 1 1 1 0

G 0 0 1 1 0 1 1 0 0

H 1 0 0 1 0 0 0 1 1

I 0 0 1 1 1 0 0 1 0

> N %\*% t(N)

Ani

Ani A B C D E F G H I

A 4 2 1 2 2 1 1 1 2

B 2 4 1 2 1 1 2 2 1

C 1 1 4 2 2 1 2 2 1

D 2 2 2 4 1 2 1 1 1

E 2 1 2 1 4 2 1 2 1

F 1 1 1 2 2 4 2 1 2

G 1 2 2 1 1 2 4 1 2

H 1 2 2 1 2 1 1 4 2

I 2 1 1 1 1 2 2 2 4

The treatment design for Phase 2 experiment can be made from the animal and treatment allocation of the Phase 1 experiment as shown below

[,1] [,2] [,3] [,4]

[1,] "controlled" "diseased" "treated" "diseased"

[2,] "controlled" "treated" "diseased" "controlled"

[3,] "treated" "controlled" "controlled" "diseased"

[4,] "controlled" "treated" "diseased" "treated"

[5,] "treated" "controlled" "controlled" "treated"

[6,] "treated" "controlled" "treated" "diseased"

[7,] "diseased" "treated" "controlled" "controlled"

[8,] "diseased" "diseased" "treated" "treated"

[9,] "diseased" "diseased" "diseased" "controlled"

The theoretical ANOVA table for this two-phase experiment with the animal and treatment allocation described can be shown as below

$ANOVA

DF e Ani Run

Between Run

Between Ani

Trt 2 1 13/15 4

Residual 6 1 49/90 4

Within

Between Ani

Trt 2 1 1931/559 0

Residual 6 1 4698/1403 0

Residual

Tag 3 1 0 0

Residual 16 1 0 0

$EF

Trt Tag eff.Trt eff.Tag

Between Run

Between Ani

Trt 45/28 15/112

Within

Between Ani

Trt 1677/164 559/656

Residual

Tag 9 1

Note this optimisation process, 0.8340019 of animal information is found in the within runs stratum, hence it is expected that treatment information may also be separated across different strata. This theoretical ANOVA table shows 133/160 of treatment information is in the within runs stratum. In addition, the coefficients of animal variance components are not identical for the treatment and the residual in the between animals within runs stratum. Hence, to test for the differences between the treatment groups, the coefficients of animal variance components needs to be adjusted.

**Example 4: assigning completely randomised design to randomised bock design**

**(Phase 1: 10 animals, 2 treatments, Phase 2: 10 MS runs and 4-plex iTRAQ tag system)**

temp 0.00826401

iteration: 10000

**Example 5: assigning completely randomised design to randomised bock design**

**(Phase 1: 12 animals, 2 treatments, Phase 2: 9 MS runs and 4-plex iTRAQ tag system)**

temp 0.00826401

iteration: 10000

Table for 4-plex system

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Phase 1 Design parameters | | Phase 2 Design parameters | | Average efficiency factors | |
| Treatments | Animals | Tags | Runs | Animal (Eliminate tag effects) | Treatment (Ignore/Eliminate tag effects) |
| 2 | 6 | 4 | 6 | 0.8936755 | 0.9166667 |
| 3 | 9 | 4 | 9 | 0.8340019 | 0.8521341 |
| 2 | 10 | 4 | 10 | 0.8231707 | 0.9 |
| 2 | 12 | 4 | 9 | 0.7833828 | 0.778/0.667 |
| 3 | 12 | 4 | 9 | 0.7833828 | 0.83125/ 0.46875 |
| 4 | 12 | 4 | 9 | 0.7833828 | 0.7935798/ 0.5732484 |
| 2 | 16 | 4 | 8 | 0.6617647 | 0.875/ 0.78125 |
| 4 | 16 | 4 | 8 | 0.6617647 | 0.7539809/ 0.5707721 |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

Table for 8-plex system

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Phase 1 Design parameters | | Phase 2 Design parameters | | Average efficiency factors | |
| Treatments | Animals | Tags | Runs | Animal (Eliminate tag effects) | Treatment (Ignore/Eliminate tag effects) |
| 2 | 6 | 8 | 6 | 0.9746386 | 0.9791667 |
| 3 | 6 | 8 | 6 | 0.9746386 | 0.9763125 |
| 2 | 10 | 8 | 5 | 0.8576644 | 0.975/ 0.1875 |
| 2 | 10 | 8 | 5 | 0.8546439 | 0.95/ 0.7281949 |
|  |  |  |  |  |  |
| 2 | 12 | 8 | 6 | 0.8597194 | 0.9583333/ 0.625 |
| 3 | 12 | 8 | 6 | 0.8597194 | 0.9601753/ 0.5846875 |
| 4 | 12 | 8 | 6 | 0.8597194 | 0.9715204/ 0.2689512 |
| 2 | 16 | 8 | 6 | 0.8076017 | 0.9375/ 0.2859926 |
| 4 | 16 | 8 | 6 | 0.8076017 | 0.9213397/0.3345515 |
| 4 | 16 | 8 | 8 | 0.8610785 | 0.9313921/0.4318623 |
| 4 | 16 | 8 | 8 | 0.8614311 | 0.909049/0.3706903 |

RBD to RBD

Cage/Ani to Run Tag