Analysis of the Ranitidine Experiment

- Model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \varepsilon$$

- Run 7 in Table 2 is dropped due to a blockage occurred in the seperation

Table 8: Least Squares Estimates, Standard Errors, t Statistics and p-values, Ranitidine Experiment (Run 7 Dropped)

Effect	Estimate	Standard Error	t	p-value
intercept	2.1850	0.5785	3.78	0.00
β_1	1.1169	0.4242	2.63	0.03
β_2	0.7926	0.4242	1.87	0.09
β_3	0.0101	0.4262	0.02	0.98
β_{11}	2.7061	0.3788	7.14	0.00
β_{22}	−0.0460	0.3786	−0.12	0.91
β_{33}	−0.1295	0.3850	−0.34	0.74
β_{12}	1.4667	0.5890	2.49	0.03
β_{13}	−0.1918	0.5890	−0.33	0.75
β_{23}	0.2028	0.5890	0.34	0.74

- only effects involving factors pH and voltage are important

NTHU STAT 5550, 2013, Lecture Notes
jointly made by Jeff Wu (GT, USA) and S. W. Cheng (NTHU, Taiwan)

Analysis of the Ranitidine Experiment

- Fitted response surface:

$$\hat{y} = 2.0373 + 1.1543 x_1 + 0.7552 x_2 + 2.7103 x_1^2 + 1.530 x_1 x_2$$

- Contour plot

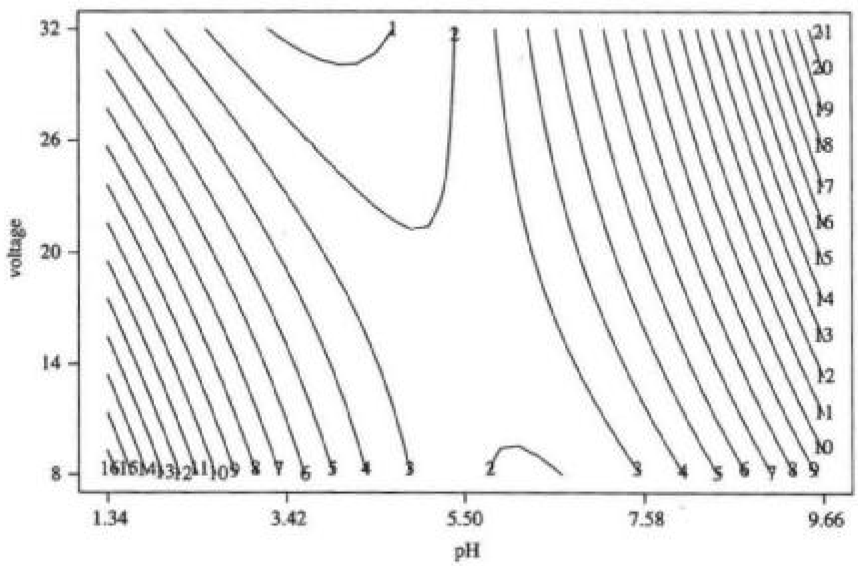

Figure 8: Estimated Response Surface, Ranitidine Experiment (Run 7 Dropped)

Analysis of the Ranitide Experiment

- A follow-up experiment in pH and voltage
  - Range of pH (A) was narrowed with levels (4.19, 4.50, 5.25, 6.00, 6.31)
  - Levels of voltage (B) were (11.5, 14.0, 20.0, 26.0, 28.5)
  - The coded values are $(-1.41, -1, 0, 1, +1.41)$

Table 9: Design Matrix and Response Data, Final Second-Order Ranitidine Experiment

	Factor		
Run order	A	B	ln CEF
2	1	−1	6.248
7	1	1	3.252
11	−1	−1	2.390
12	−1	1	2.066
3	−1.41	0	2.100
8	1.41	0	9.445
9	0	1.41	1.781
1	0	−1.41	6.943
4	0	0	2.034
5	0	0	2.009
6	0	0	2.022
10	0	0	1.925
13	0	0	2.113

NTHU STAT 5550, 2013, Lecture Notes
jointly made by Jeff Wu (GT, USA) and S. W. Cheng (NTHU, Taiwan)

Analysis of Final Ranitidine Experiment

- Model: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon$
- Analysis result:

Table 10: Least Squares Estimates, Standard Errors, t Statistics and p-values, Final Second-Order Ranitidine Experiment

Effect	Estimate	Standard Error	t	p-value
intercept	2.0244	0.5524	3.66	0.0080
β_1	1.9308	0.4374	4.41	0.0031
β_2	−1.3288	0.4374	−3.04	0.0189
β_{12}	−0.6680	0.6176	−1.08	0.3153
β_{11}	1.4838	0.4703	3.15	0.0160
β_{22}	0.7743	0.4703	1.65	0.1437

Analysis of the Ranitidine Experiment

- Fittied model:

$$\hat{y} = 2.0244 + 1.9308x_1 - 1.3288x_2 - 0.6680x_1x_2 + 1.4838x_1^2 + 0.7743x_2^2$$
$$= \hat{\beta}_0 + \mathbf{x}^T\mathbf{b} + \mathbf{x}^T\mathbf{B}\mathbf{x},$$

where
$$\mathbf{b} = (1.9308, -1.3288)^T$$

and
$$\mathbf{B} = \begin{bmatrix} 1.4838 & -0.3340 \\ -0.3340 & 0.7743 \end{bmatrix}.$$

- The stationary point is $\mathbf{x}_s = -\frac{1}{2}\mathbf{B}^{-1}\mathbf{b} = (-0.5067, 0.6395)^T$, which yields $y_s = 1.1104$
- The eigen-decomposition of $\mathbf{B}$ yields eigenvalues

$$\Lambda = diag(1.6163, 0.6418),$$

and eigenvectors
$$\mathbf{P} = \begin{bmatrix} -0.9295 & -0.3687 \\ 0.3687 & -0.9295 \end{bmatrix}.$$

NTHU STAT 5550, 2013, Lecture Notes
jointly made by Jeff Wu (GT, USA) and S. W. Cheng (NTHU, Taiwan)

Analysis of the Ranitidine Experiment

- Since both λ_1 and λ_2 are positive, $y_s = 1.1104$ is the minimum value which is achieved at $\mathbf{x}_s$ (pH of 4.87 and voltage of 23.84).

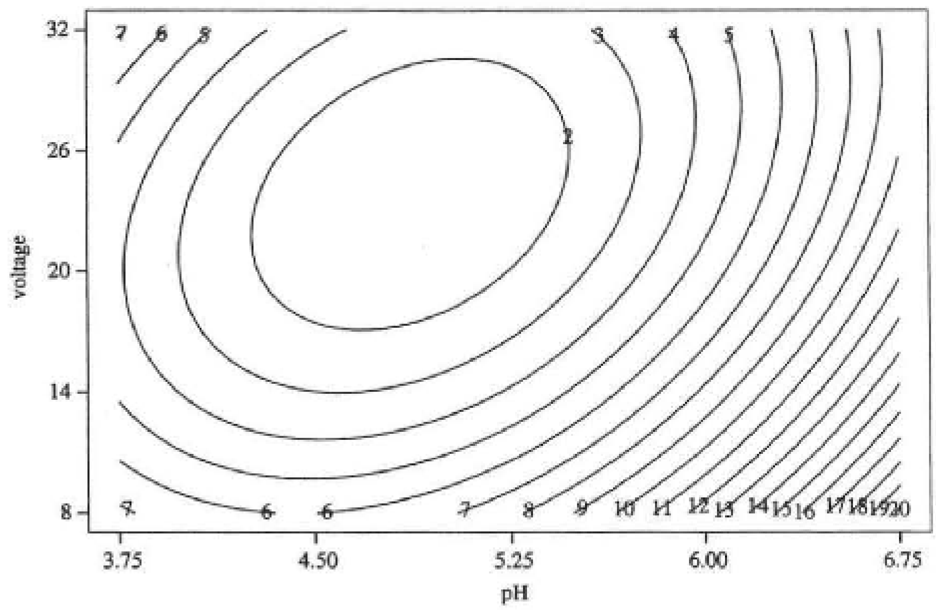

Figure 9: Estimated Response Surface, Final Second-Order Ranitidine Experiment

- **Reading**: textbook, 10.5

Central Composite Designs

- The k input factors in coded form are denoted by $\mathbf{x} = (x_1, \ldots, x_k)$.
- A second-order model has $1 + k + k + \binom{k}{2} = \frac{(k+1)(k+2)}{2}$ parameters
- A central composite design consists of the following three parts:
  1. n_f *cube points* (or *corner points*) with $x_i = -1$ or 1 for $i = 1, \ldots, k$. They form the *factorial portion* of the design.
  2. n_c *center points* with $x_i = 0$ for $i = 1, \ldots, k$.
  3. $2k$ *star points* (or *axial points*) of the form $(0, \cdots, x_i, \cdots, 0)$ with $x_i = \alpha$ or $-\alpha$ for $i = 1, \ldots, k$.
  - For the ranitidine experiment, the cube points are the 2^3 design, $n_c = 6$ and $\alpha = 1.66$.
  - $N = n_f + 2k + 1 \geq \frac{(k+1)(k+2)}{2} \Rightarrow n_f \geq \frac{k(k-1)}{2}$
- The central composite design can be used in a *single* experiment or in a *sequential* experiment.

NTHU STAT 5550, 2013, Lecture Notes
jointly made by Jeff Wu (GT, USA) and S. W. Cheng (NTHU, Taiwan)

Central Composite Designs

- In choosing a central composite design, there are three issues:
  1. choosing the factorial portion of the design,
  2. determining the number of center points,
  3. determining the α value for the star points.

Central Composite Designs - Cube Points

- Function of the three parts in fitting a second-order model:
  - cube points: estimating linear main effects and interactions
  - center points: estimating overall quadratic main effects and $\hat{\sigma}$ (replicates)
  - star points: estimating and dealiasing linear and quadratic main effects
- **Theorem**. In any central composite design whose factorial portion is a 2^{k-p} design that does not use any main effect as a defining relation, the following parameters in (2) are estimable: β_0, β_i, β_{ii}, $i = 1, \ldots, k$, and one β_{ij} selected from each set of aliased effects for $i < j$. It is not possible to estimate more than one β_{ij} from each set of aliased effects.

NTHU STAT 5550, 2013, Lecture Notes
jointly made by Jeff Wu (GT, USA) and S. W. Cheng (NTHU, Taiwan)

Central Composite Designs - Cube Points

- It is interesting to note that
  - even defining words of length two (for $k = 2$ case) are allowed and
  - words of length four are worse than words of length three.
- Any resolution III design whose defining relation does not contain words of length four is said to have **resolution** III*.
- *Any central composite design whose factorial portion has resolution III* is a second-order design.*
- For the estimability of the parameters in the second-order model, one can only use the cube and star points of the central composite design if $\alpha \neq \sqrt{k}$. Such a design is referred to as a *composite design* and its run size is $n_f + 2k$.
  - the smallest designs without center points in the Table 11 for k= 2, 3, 5, 6 and 7 have the minimal run size and are saturated.

Central Composite Designs - Cube Points

Table 11: Central Composite Designs for $2 \leq k \leq 7$

k	$(k+1)(k+2)/2$	N	n_f	Factorial Portion (cube points)
2	6	7	2	$2^{2-1}(\mathbf{I}=AB)$
2	6	9	4	2^2
3	10	11	4	$2_{III}^{3-1}(\mathbf{I}=ABC)$
3	10	15	8	2^3
4	15	17	8	$2_{III}^{4-1}(\mathbf{I}=ABD)$
4	15	20	11	11×4 submatrix of 12−run PB design
4	15	25	16	2^4
5	21	22	11	11×5 submatrix of 12−run PB design
5	21	23	12	12×5 submatrix of 12−run PB design
5	21	27	16	$2_{V}^{5-1}(\mathbf{I}=ABCDE)$
6	28	29	16	$2_{III^*}^{6-2}(\mathbf{I}=ABE=CDF=ABCDEF)$
7	36	37	22	22×7 submatrix given in Table 10A.2 (textbook)
7	36	38	23	23×7 submatrix given in Table 10A.3 (textbook)
7	36	47	32	$2_{III^*}^{7-2}(\mathbf{I}=ABCDF=DEG)$

NTHU STAT 5550, 2013, Lecture Notes
jointly made by Jeff Wu (GT, USA) and S. W. Cheng (NTHU, Taiwan)

Central Composite Designs - Axial Points

- The efficiency of the parameter estimates is increased by pushing the axial points toward the extreme.
- In general, α should be chosen between 1 and $\sqrt{k}$ and rarely outside this range.
- For α=1, the axial points are placed at the center of the faces of the cube.
  - The design is therefore called the *face center cube*.
  - They are the only central composite designs that require three levels.
  - They are effective designs if the design region is a cube.
- For $\alpha=\sqrt{k}$, the axial points and cube points lie on the same sphere.
  - The design is often referred to as a *spherical design*.
  - They are effective designs if the design region is spherical.
  - For large k, this choice should be taken with caution.
- In general the choice of α depends on the geometric nature of and the practical constraints on the design region.

Central Composite Designs - Axial Points

- A design is called **rotatable** if $Var(\hat{y}(\mathbf{x}))$ depends only on

$$\|\mathbf{x}\| = (x_1^2 + \cdots + x_k^2)^{1/2},$$

that is, if the accuracy of prediction of the response is the same on a sphere around the center of the design.

- For a central composite design whose factorial portion is a 2^{k-p} design of resolution V, it can be shown that rotatability is guaranteed by choosing

$$\alpha = \sqrt[4]{n_f}. \quad (5)$$

- Eqn. (5) serves as a useful guide even when the factorial portion does not have resolution V.

NTHU STAT 5550, 2013, Lecture Notes
jointly made by Jeff Wu (GT, USA) and S. W. Cheng (NTHU, Taiwan)

Central Composite Designs - Axial Points

- The rotatability criterion should not, however, be strictly followed.
  - Its definition depends on how the x_i variables are coded. This lack of invariance with respect to the coding scheme makes it less attractive as a theoretical criterion.
  - Because rotatability often requires that the factorial portion has resolution V, use of a smaller composite design whose factorial portion has a lower resolution has not been popular. However, some of the central composite designs whose factorial portions have resolution V or higher have excessive degrees of freedom.
    - For $k = 6$, the smallest design that satisfies the rotatability requirement is a 2_{VI}^{6-1} design for the factorial portion.
    - The total run size of the central composite design is 45, which has 17 additional degrees of freedom over the minimal run size.
- **Near rotatability** will be a more reasonable and practical criterion.

Central Composite Designs - Center Points

- The center points can
  - provide information on the pure error and
  - help stabilize the variance of the predicted response $\hat{y}$.
- For $\alpha=\sqrt{k}$, at least one center point is required for the estimability of the parameters in the second-order model. (Why?) Otherwise the variance of $\hat{y}(\mathbf{x})$ becomes infinite.
- To stabilize the prediction variance, the rule of thumb suggests
  - 3 to 5 runs at the center point when $\alpha \approx \sqrt{k}$,
  - 1 or 2 runs at the center point when $\alpha \approx 1$
  - 2 to 4 runs should be considered between these two extremes
- To estimate the error variance, more than 4 or 5 runs may be required.

v **Reading**: textbook, 10.7

NTHU STAT 5550, 2013, Lecture Notes
jointly made by Jeff Wu (GT, USA) and S. W. Cheng (NTHU, Taiwan)

Box-Behnken Designs

- A family of three-level second-order designs
- Combining two-level factorial designs with balanced (or partially balanced) incomplete block designs in a particular manner
- Illustration: A BIBD with 3 treatments and 3 blocks of size 2

Block	Treatment 1	Treatment 2	Treatment 3
1	×	×	
2	×		×
3		×	×

	Factors x_1	Factors x_2	Factors x_3
1	±1	±1	0
2	±1	0	±1
3	0	±1	±1
4	0	0	0

  - run size $= 3 \times 2^2 + n_c$
  - Why is it a second-order design?