

1. (a) One would usually want to measure covariates before assigning treatment so that treatments could be assigned on the basis of covariates. By assigning each treatment to a variety of values of covariates (for instance, each unit in a block receiving different treatment), it becomes less likely the effect of the covariate is confounded with the effect of the treatment.
- (b) Let v be an additional unit such that, where S is the set of all units, $\forall v' \in S \setminus \{v\}, R(v, v') \neq 1$. Since we can not apply the given logic to v , it need not be the case $R(v, v) = 1$, so the relation is not necessarily reflexive.
- (c) Site, race, sex, and marital status are unordered qualitative factors, since they each take on a finite number of levels which admit not obvious ordering. I would classify age as a quantitative covariate, since it can be measured as a continuous (as opposed to discrete) value; in practice however, it can also be treated as an ordered qualitative factor, since it is usually recorded in discrete year values.
- (d) (i) For the $y \sim A + \mathbf{1}$ model, the model matrix is

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

The second column, corresponding to level 1 of A, would be omitted under the default in R (treatment constraint). For the $y \sim A - \mathbf{1}$ model, the model matrix is

$$\begin{bmatrix} -1 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{bmatrix}$$

Once again, the second column would be omitted. These two models are equivalent in the sense they would yield the same predictions. This is because they define the same subspace since $-\mathbf{1} \in \text{span}(\mathbf{1})$. For $y \sim A + B$, the model matrix would be

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Here, the 4th column, corresponding to the 1 level of B would be omitted, since it is in the span of the rest of the columns.

- (ii) If we treat A as a quantitative covariate, then for the $y \sim A + \mathbf{1}$ model, the model matrix is

$$\begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 1 & 2 \\ 1 & 3 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

With no columns omitted in R. For the $y \sim A - \mathbf{1}$ model, the model matrix is

$$\begin{bmatrix} -1 & 3 \\ -1 & 2 \\ -1 & 2 \\ -1 & 3 \\ -1 & 1 \\ -1 & 2 \\ -1 & 3 \\ -1 & 1 \\ -1 & 1 \end{bmatrix}$$

With no columns omitted in R either. This model still defines an equivalent linear space to the $y \sim A + \mathbf{1}$ model, so would yield identical predictions. Finally, For $y \sim A + B$, the model matrix would be

$$\begin{bmatrix} 3 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 2 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

Which would have no columns omitted.

- (iii) For each of these subspaces, the dimension is the same as the dimension of the model matrix as created above. The squared length of the orthogonal projection is, when the model matrix is X , $\|X(X'X)^{-1}X'y\|^2$ and the squared length of the orthogonal complement is $\|(I - X(X'X)^{-1}X')y\|^2$. This yields:

	Dimension	Sqrd Len Proj	Sqrd Len Comp
1	1	258.89	85.63
A	3	330.52	13.99
B	3	262.84	81.67
A+B	5	332.59	11.92
A*B	7	343.84	0.67

- (iv) I would use the sum of squares from A. Specifically, I would compare the marginal difference in sum of squares between the projection onto the A space and onto the $\mathbf{1}$ space, and compare this to the sum of squares in the complement space to A. It appears B has little if any effect, so it can probably be safely ignored for this purpose. If there is no treatment effect, these two values should

both be χ^2 variables with the respective degrees of freedom, so we can use them to construct a F test:

$$\frac{\frac{330.52-258.89}{3-1}}{\frac{13.99}{6}} = 15.361$$

Checking against a $F_{2,6}$ distribution, this corresponds to a p-value of .0044.

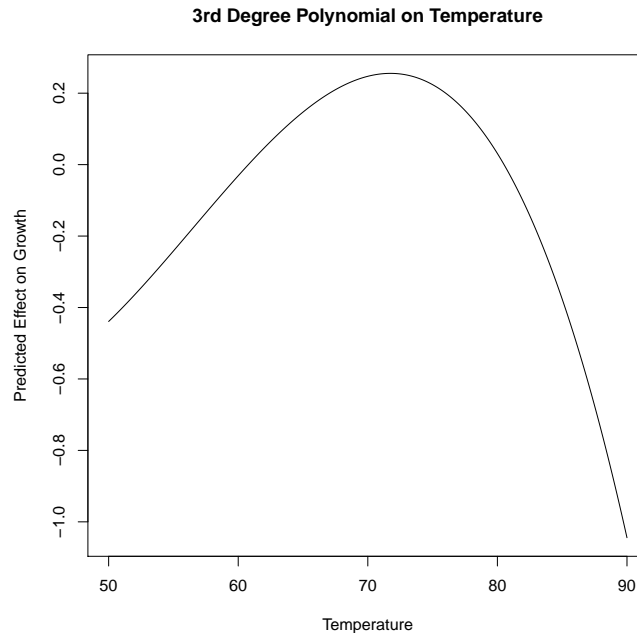
2. (i) This table can be gotten directly from the anova command in R

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Isolate	7	0.3481	0.0497	4.2338	0.0013
Temp	6	2.7234	0.4539	38.6397	0.0000
Resid	42	0.4934	0.0117		

- (ii) This table is a bit more complicated. I generated most of the output from a series of anova commands (Listing the spaces in increasing size so as to get the right quotient spaces), but then recalculated the F and p-values based on the overall residuals from the first model:

	Df	Sum Sq	Mean Sq	F	pvalue
P1/1	1	0.0117	0.0117	0.9974	0.3237
P2/P1	1	2.5678	2.5678	218.5976	0.0000
P3/P2	1	0.0999	0.0999	8.5061	0.0057
F value	1	0.0134	0.0134	1.1446	0.2908
Pr(>F)	2	0.0305	0.0152	1.2963	0.2843
Temp/P3	3	0.0439	0.0146	1.2457	0.3053
Total Temp	6	2.7234	0.4539	38.6397	0.0000
Resid	42	0.4934	0.0117		

- (iii) This table strongly indicates that temperature an effect on fungal growth, since temperature as a categorical variable has a very low p-value on it. However, lower order polynomials seem to be sufficient to capture the effect; while the linear fit is poor, adding both a 2nd and 3rd degree polynomial seems to cause statistically significant improved fits. After this though, greater complexity does not seem to cause statistically significant improvement, indicating a 3rd degree polynomial is optimal.
- (iv) I've plotted the cubic fit below:



Based off this fit, one would predict growth rate will be maximized by a temperature of approximately 71.75.

3. (i) For each conservative matrix V , it must be the case $\forall i, j$ that $V(i_0, i_1) + V(i_1, i_0) = 0 \implies V(i_0, i_1) = -V(i_1, i_0)$. This means that V is skew-symmetric. On the other hand, for

$$W = \begin{bmatrix} 0 & -4 & -1 \\ 4 & 0 & -2 \\ 1 & 2 & 0 \end{bmatrix}$$

We have that $W(3, 1) + W(1, 2) + W(2, 3) = 1 - 4 - 2 \neq 0$, so not all skew-symmetric matrices are conservative.

If a matrix V is additive, for a path of length m where $i_m = i_0$,

$$\sum_{j=0}^{m-1} V(i_j, i_{j+1}) = \alpha_{i_0} + \sum_{j=1}^{m-1} (-\alpha_{i_j} + \alpha_{i_j}) - \alpha_{i_0} = 0$$

So V is conservative. We may also conclude the converse by the following logic on any conservative matrix V :

$$\begin{aligned} 0 &= V(u_1, u_2) + V(u_2, u_3) + V(u_3, u_4) + V(u_4, u_1) \\ \implies -V(u_4, u_1) - V(u_2, u_3) &= V(u_1, u_2) + V(u_3, u_4) \\ \implies V(u_1, u_4) + V(u_3, u_2) &= V(u_1, u_2) + V(u_3, u_4) \end{aligned}$$

This is only possible if, for all u_i, u_j , $V(u_i, u_j) = f(u_i) + g(u_j)$ for some functions f, g . Since

$$\begin{aligned} V(u_1, u_3) &= V(u_1, u_2) + V(u_2, u_3) \\ f(u_1) + g(u_3) &= f(u_1) + g(u_2) + f(u_2) + g(u_3) \\ f(u_2) &= -g(u_2) \end{aligned}$$

Proving that $V(u_i, u_j) = f(u_i) - f(u_j)$ and thus that V is additive.

Finally, the space of 6×6 conservative matrices had dimension 5. This is because the matrix is minimally defined by $\{V(1, 2), V(1, 3), V(1, 4), V(1, 5), V(1, 6)\}$. All the remaining entries can be calculated as a linear combination of these five values: $V(u, j) = V(u, 1) + V(1, j) = -V(1, u) + V(1, j)$. No fewer values will suffice for this purpose.

- ii We can set this up as a typical least squares problem by transforming the $k \times k$ matrix Y into a $k^2 \times 1$ vector y . Then, if R is the usual matrix of 1s and 0s for the row as a categorical variable, and C is the same for the column, our model matrix is $X = R - C$. This gives us $E(y) = X\alpha$. After omitting one of the columns of X due to collinearity, we can then produce the normal least squares estimate $\alpha = (X'X)^{-1}X'y$. $\alpha_i - \alpha_j = (\alpha_i + c) - (\alpha_j + c)$, so $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5\}$ and $\{\alpha_1 + c, \alpha_2 + c, \alpha_3 + c, \alpha_4 + c, \alpha_5 + c\}$ are equivalent parameters, demonstrating that one parameter needs to be omitted.

When we preform this procedure on the data, we get these predictions for the α by electrolyte:

	Zn	Fe	Pb	Cu
O	-0.3908	-0.8408	-0.9138	-1.2426
A	-0.3962	-0.8690	-0.9624	-1.3284
K	-0.4398	-0.8850	-0.9566	-1.3106

As noted, the fifth value is a free parameter and was omitted.

- iii These two models are one just including the α , and one where the α are interacted with electrolyte. Running these two models and comparing them with ANOVA, we get:

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	70	0.16				
2	62	0.13	8	0.03	1.99	0.0623

By the interaction, we don't see quite enough of an improvement to push past the .05 p-value threshold, though it is a borderline case.

- iv Our data set only has factor covariates, so no monotone transformation will have an effect on the predictions, and non-monotone transformation wouldn't make sense. If there was evidence of a lack of normality or non-constant variance, a transformation might improve our statistical inference, but checking various residual plots leaves little indication that this is necessary.

5. This can be formulated as a linear model after transformation:

$$\log(Y) \sim \alpha - \beta_T t$$

If we are trying to model the expected half-life h for a given sample under this model, we can formulate the problem as

$$\begin{aligned} \log(.5) &= \hat{\beta}_T h \\ \frac{1}{h} &= \frac{\hat{\beta}_T}{\log(.5)} \end{aligned}$$

The value on the right is t-distributed under the assumptions of a normal model, so we can generate a 95% confidence interval by $\hat{\beta}_T \pm t_{n-p}^{0.25} se(\hat{\beta})$. Then, we can take the inverse of each boundary to get a confidence interval on the h . We must keep in mind though that any value that is 0 or negative before taking the inverse corresponds to a lack of decay and asorbic acid levels not actually falling. This yields the following confidence intervals for half-life (in weeks):

<i>Temperature</i>	<i>LowerBound</i>	<i>Estimate</i>	<i>UpperBound</i>
0	67.63	948.55	∞
10	20.70	28.92	47.99
20	4.89	5.24	5.65

It is worth noting that this does not include any prediction error. However, that error's contribution would be relatively small. This model accounts for the variance among replicates; that variance contributes to the variance in the mean which we are fitting.