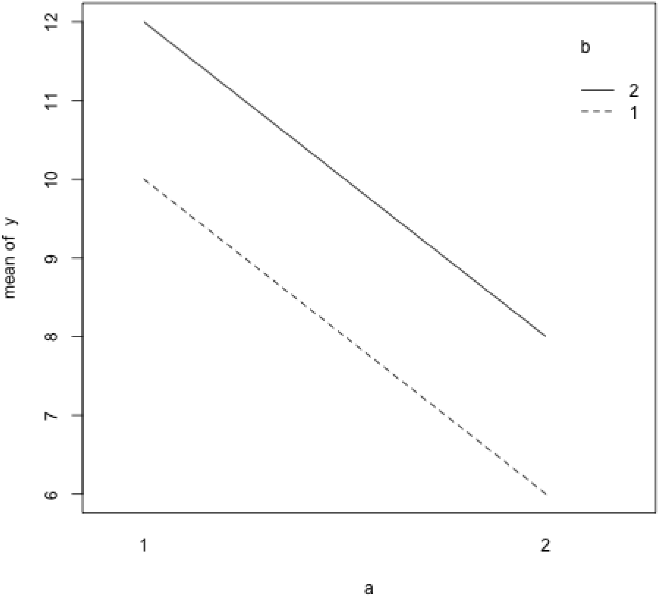
Two-way ANOVA

Notation: Factor A has a levels, factor B has b levels.
Factorial design: treatments represent all combinations of levels of A and B.
$y_{ijk} = \mu_{ij} + \epsilon_{ijk}, i = 1, \cdots, a; j = 1, \cdots, b; k = 1, \cdots, n.$

Suppose $\mu_{11} = 10, \mu_{12} = 12, \mu_{21} = 6, \mu_{22} = 8.$
$\mu_{1\cdot} = (10 + 12)/2 = 11, \mu_{2\cdot} = 7, \mu_{\cdot\cdot} = 9.$

True main effect of A_1 is $\alpha_1 = \mu_{1\cdot} - \mu_{\cdot\cdot} = 11 - 9 = 2,$
similarly, $\alpha_2 = -2.$

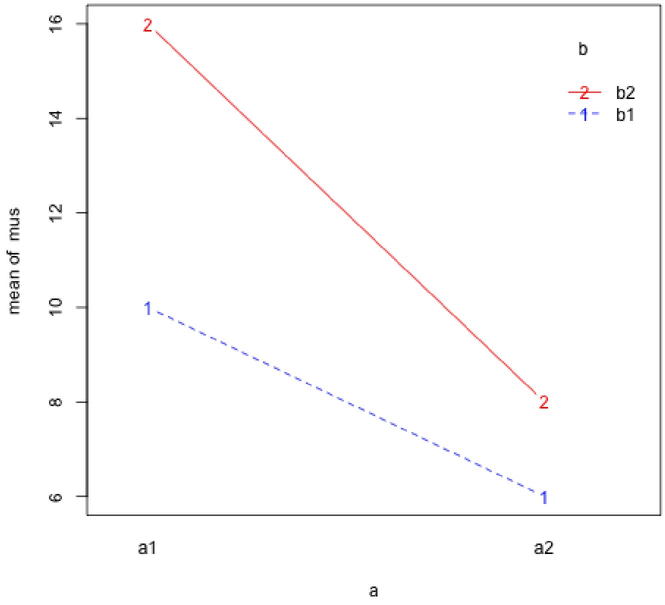
Also we can get true main effect for B: $\beta_1 = -1, \beta_2 = 1.$


b
2
1
mean of y
12
11
10
9
8
7
6
1
2
a

Interaction

Each $\mu_{ij} = \mu_{\cdot\cdot} + \alpha_i + \beta_j$,
and $y_{ijk} = \mu_{\cdot\cdot} + \alpha_i + \beta_j + \epsilon_{ijk}$.
This is the no interaction model.

Interaction model:
$\mu_{11} = 10, \mu_{12} = 16, \mu_{21} = 6, \mu_{22} = 8$.
The effect of one factor depends on the level of the other factor.
$\mu_{ij} = \mu_{\cdot\cdot} + \alpha_i + \beta_j + \alpha\beta_{ij}$
and $y_{ijk} = \mu_{\cdot\cdot} + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk}$.
where $\alpha\beta_{ij} = \mu_{ij} - (\mu_{\cdot\cdot} + \alpha_i + \beta_j)$.



16
14
12
10
8
6
mean of mus
b
b2
b1
a1
a2
a

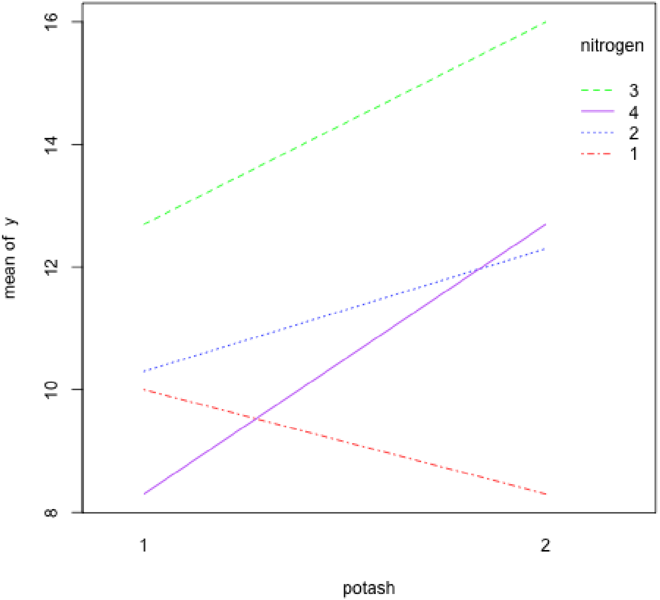
estimate the main and interaction effects

$$\hat{\alpha}_i = \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot}$$
$$\hat{\beta}_j = \bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot\cdot\cdot}$$
$$\widehat{\alpha\beta}_{ij} = \bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\cdot\cdot\cdot}$$

		Nitrogen			
		5%	10%	15%	20%
Potash	10%	10.0	10.3	12.7	8.3
	20%	8.3	12.3	16.0	12.7

MSE= 3.625.

a. Find $\hat{\alpha}_i, i = 1, 2$.
b. Find $\hat{\beta}_j, j = 1, 2, 3, 4$.
c. Find $\widehat{\alpha\beta}_{ij}, i = 1, 2; j = 1, 2, 3, 4$.



16
14
12
10
8
mean of y
nitrogen
3
4
2
1
1
2
potash

$\bar{y}_{1..} = 10.325, \bar{y}_{2..} = 12.325,$
$\bar{y}_{...} = 11.325,$
Hence the estimated Potash effects are:
$\hat{\alpha}_1 = 10.325 - 11.325 = -1, \hat{\alpha}_2 = 1.$

similarly,
$\bar{y}_{.1.} = 9.15, \bar{y}_{.2.} = 11.3,$
$\bar{y}_{.3.} = 14.35, \bar{y}_{.4.} = 10.5,$
so the estimated Nitrogen effects are
$\hat{\beta}_1 = -2.175, \hat{\beta}_2 = -0.025, \hat{\beta}_3 = 3.025, \hat{\beta}_4 = -0.825.$

The estimated interaction effects:
$\widehat{\alpha\beta}_{11} = 10.0 - 10.325 - 9.15 + 11.325 = 1.85.$

interaction plot

```
y11=c(8,8,9,12);y12=c(11,11,12,13)
y21=c(5,6,6,6); y22=c(7,8,8,9)
y= c(y11, y12, y21, y22)
a = c(rep(1,8), rep(2,8))
b= c(rep(1,4), rep(2,4), rep(1, 4), rep(2, 4))
a=factor(a)
b=factor(b)
interaction.plot(a,b,y)
output = lm (y~ a+b+a*b)
anova(output)
```

```
                  B
           1             2
-----|------------|--------
A 1  |  8,8,9,12  |11,11,12,13      mean:10.5
-----|------------|---------
  2  |  5,6,6,6   |7,8,8,9          mean: 6.875

                             grand mean: 8.6875
```

Compute *SSA*.
$\bar{y}_{1\cdot\cdot} = 10.5, \bar{y}_{2\cdot\cdot} = 6.875, \bar{y}_{\cdot\cdot\cdot} = 8.6875.$
$b = 2, n = 4,$
SSA$= 2 * 4 * [(10.5 - 8.6875)^2 + (6.875 - 8.6875)^2] = 52.56$

ANOVA Table

$$
\begin{aligned}
SST_c &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{\cdots})^2,\\
SSA &= bn\sum_{i=1}^{a}(\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdots})^2,\\
SSB &= an\sum_{j=1}^{b}(\bar{y}_{\cdot j\cdot}-\bar{y}_{\cdots})^2,\\
SSAB &= n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot j\cdot}+\bar{y}_{\cdots})^2,\\
SSE &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{ij\cdot})^2.
\end{aligned}
$$

Fact: $SST_c = SSA + SSB + SSAB + SSE$.

Source of Variation	df	SS	MS	F	P-value
A	a-1	SSA	MSA	MSA/MSE	
B	b-1	SSB	MSB	MSB/MSE	
A*B	(a-1)*(b-1)	SSAB	MSAB	MSAB/MSE	
Error	ab(n-1)	SSE	MSE		
Total	N-1	SST_c			

$SST_c = SSA + SSB + SSAB + SSE$
$N - 1 = (a - 1) + (b - 1) + (a - 1)(b - 1) + ab(n - 1)$
where $N = abn$.

F test

Factor A effect:
$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$
$F = \frac{MSA}{MSE}$,

Factor B effect: $F = \frac{MSB}{MSE}$,
$H_0 : \beta_1 = \beta_2 = \cdots = \beta_b = 0$

AB interaction effect: $F = \frac{MSAB}{MSE}$.
$H_0 : \alpha\beta_{11} = \cdots = \alpha\beta_{ab} = 0$

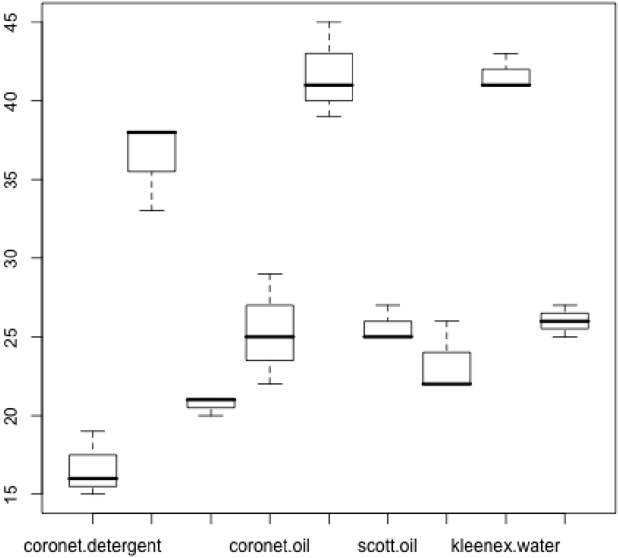
Paper Towel: Amount of liquid absorbed (mL)

	Water	Detergent	Oil	
Coronet	26,22,22	19,16,15	22,25,29	mean:21.78
Kleenex	43,41,41	33,38,38	39,41,45	mean:39.89
Scott	27,26,25	21,20,21	27,25,25	

Paper towel example

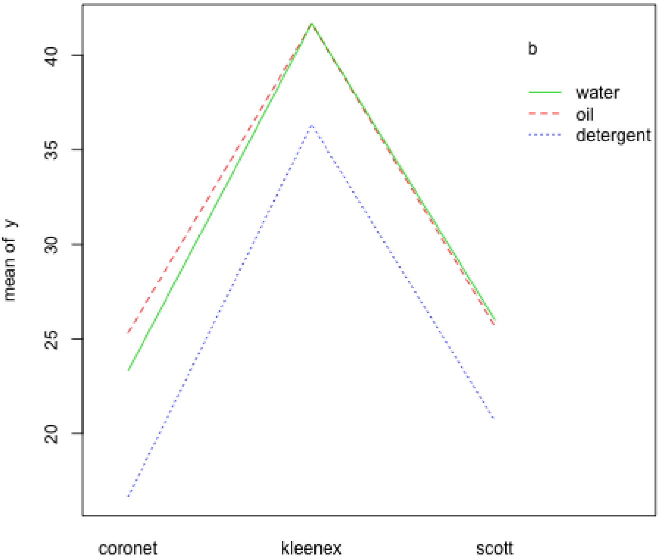
```
y = c(26,22,22,19,16,15,22,25,29,43,41,41,33,38,38,39,41,45,
27,26,25,21,20,21,27,25,25)
a = c(rep("coronet",9),rep("kleenex",9),rep("scott",9))
b1 =c(rep("water",3),rep("detergent",3),rep("oil",3))
b = c(b1,b1,b1)
a =factor(a)
b=factor(b)
interaction.plot(a,b,y)
out = lm (y~a+b+a*b)
anova(out)
boxplot(y~a+b)
output = aov (y~a+b)
TukeyHSD(output)
```

Boxplot



45
40
35
30
25
20
15
coronet.detergent
coronet.oil
scott.oil
kleenex.water

Interaction plot



```
> output <- aov(y~a+b+a*b)
> summary(output)
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq  F value    Pr(>F)
a          2 1747.19  873.59 180.0534 1.256e-12 ***
b          2  221.41  110.70  22.8168 1.160e-05 ***
a:b        4   12.59    3.15   0.6489     0.635
Residuals 18   87.33    4.85
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

```
> output <- aov(y~a+b)
> summary(output)

Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)
a          2 1747.19  873.59 192.333 1.162e-14 ***
b          2  221.41  110.70  24.373 2.630e-06 ***
Residuals 22   99.93    4.54
```

```
> TukeyHSD(output)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = y ~ a + b)
$a
                      diff         lwr        upr     p adj
kleenex-coronet 18.111111  15.5873279  20.634894 0.0000000
scott-coronet    2.333333  -0.1904499   4.857117 0.0734828
scott-kleenex  -15.777778 -18.3015610 -13.253995 0.0000000

$b
                       diff       lwr      upr     p adj
oil-detergent     6.3333333  3.809550 8.857117 0.0000070
water-detergent   5.7777778  3.253995 8.301561 0.0000252
water-oil        -0.5555556 -3.079339 1.968228 0.8460364
```

check: qtukey(0.95,3,22)/sqrt(2)=2.512,
$39.89 - 21.78 \pm 2.512 * \sqrt{4.54} * \sqrt{1/9 + 1/9} = 18.11 \pm 2.52 =$
$(15.59, 20.63)$.

Problem 6.2

```
shoot =
read.table("http://educ.jmu.edu/ chen3lx/math321/shoot.txt",header=T
interaction.plot(shoot$hand, shoot$distance, shoot$y)
out =lm (y~distance+hand+distance*hand,shoot)
anova(out)
output = aov (shoot$y ~ shoot$distance)
TukeyHSD(output)
```

prob 6.4.

The anova table shows there is significant interaction effect.

```
>  butter <- read.table("/Users/lchen/Sites/math321/
   butter.txt",header=T)
> out <- lm(y~brand*cookmethod,data=butter)
> anova(out)
Analysis of Variance Table

Response: y
                 Df  Sum Sq Mean Sq F value    Pr(>F)
brand             2  2683.0  1341.5  6.0931   0.01492 *
cookmethod        1 11806.7 11806.7 53.6263 9.185e-06 ***
brand:cookmethod  2  1470.8   735.4  3.3401   0.07027 .
Residuals        12  2642.0   220.2
```

Compare the effect of cookmethod (stove vs oven) fixing brand =lakes.

Compare the effect of brand (lakes vs value, lakes vs cabot, value vs cabot) fixing cookmethod = stove.

```
y = butter$y
brand = butter $ brand
cookmethod =butter $ cookmethod
y11= y[brand=="lakes" & cookmethod=="stove"]
mean(y11)
```

The means are:

brand/cookmethod	stove	oven
lakes	152.00	182.67
cabot	111.00	185.67
value	153.67	202.00

oven vs stove at brand level lakes:
$182.67 - 152.00 \pm 2.179 * \sqrt{220.2} * \sqrt{1/3 + 1/3} =$
$30.67 \pm 26.40 = (4.27, 57.07)$.
$R : qt(0.975, 12)$ noting $m = 1$ here.

Compare the effect of brand fixing cookmethod = stove.
$m = 3$,
critical value: qtukey(0.95,3,12)/sqrt(2)=2.668.
margin of error $= 2.668 * \sqrt{220.2} * \sqrt{1/3 + 1/3} = 32.32$.
lakes - cabot: $152.00 - 111.00 \pm 32.32 = (8.68, 73.32)$
lakes - value: $152 - 153.67 \pm 32.32 = (-33.99, 30.65)$.
cabot - value: $111.00 - 153.67 \pm 32.32 = (-74.99, -10.35)$.

```
 install.packges("emmeans")
 library(emmeans)
 out=lm(y~brand*cookmethod,data=butter)
confint(emmeans(out,pairwise~brand|cookmethod),level=0.95)
cookmethod = oven:
 contrast        estimate       SE df  lower.CL   upper.CL
 cabot - lakes   3.000000 12.11519 12 -29.32167  35.321669
 cabot - value -16.333333 12.11519 12 -48.65500  15.988336
 lakes - value -19.333333 12.11519 12 -51.65500  12.988336

cookmethod = stove:
 contrast        estimate       SE df  lower.CL   upper.CL
 cabot - lakes -41.000000 12.11519 12 -73.32167  -8.678331
 cabot - value -42.666667 12.11519 12 -74.98834 -10.344997
 lakes - value  -1.666667 12.11519 12 -33.98834  30.655003

Confidence level used: 0.95
Conf-level adjustment: tukey method for comparing a family
of 3 estimates
```