

I. Excel 상관분석

II. R 상관분석

1. 공분산

- 두 변수가 함께 변하는 정도를 측정

$$\text{Cov}(X,Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n-1}$$

- 공분산의 부호와 크기로 두 변수의 관계를 판단

- 부호가 양(+) : 두 변수가 같은 방향으로 변화 즉, X변수가 증가(감소)하면 Y변수도 증가(감소)

예를 들면, 통화량 증가율과 물가 상승률

크기가 클수록 두 변수는 강한 양(+)의 관계를 가짐

- 부호가 음(-) : 두 변수가 반대 방향으로 변화 즉, X변수가 증가(감소)하면 Y변수는 감소(증가)

예를 들면, 경제성장률과 실업률

크기가 작을수록 두 변수는 강한 음(-)의 관계를 가짐

- 크기가 0 : 두 변수는 관계가 없음

- 두 변수의 측정단위에 따라 공분산의 부호는 변하지 않지만 공분산의 크기는 달라짐

- 예를 들어 X의 측정단위 변화로 X의 값이 원래보다 10배 증가하면 공분산의 크기도 10배 증가

- X 및 Y의 측정단위 변화로 X 및 Y의 값이 각각 원래보다 10배씩 증가하면 공분산의 크기는 100배 증가

- 공분산의 크기는 측정단위에 영향을 받기 때문에 동일한 집단이라도 측정단위에 따라 달라지는 한계로 많이 이용되지 않음

2. 상관계수

- 상관분석은 독립변수와 종속변수의 구분이 없는 두 확률변수 간의 선형성의 정도를 측정한 상관계수를 추정하고 검정하는 것을 말함
- 상관계수는 두 변수 간 선형관계의 밀접도를 측정하는 통계량

$$\text{Corr}(X,Y) = \rho_{X,Y} = \frac{\text{Cov}(X,Y)}{S_X S_Y}$$

단, S_X, S_Y 는 X 및 Y의 표준편차를 각각 나타냄

- 상관계수의 부호와 크기로 두 변수의 관계를 판단
 - 부호가 양(+) : 두 변수들이 서로 같은 방향으로 변화 즉, X변수가 증가(감소)하면 Y변수도 증가(감소)
 - 부호가 음(-) : 두 변수들이 서로 반대 방향으로 변화 즉, X변수가 증가(감소)하면 Y변수는 감소(증가)
 - 크기가 0 : 두 변수는 관계가 없음
- 상관계수는 -1과 1사이의 값을 가짐
 - 상관계수의 크기가 양(+)이면서 1에 가까워질수록 두 변수는 강한 양(+)의 관계를 가짐
 - 상관계수의 크기가 음(-)이면서 -1에 가까워질수록 두 변수는 강한 음(-)의 관계를 가짐
 - 상관계수의 크기가 0이면 두 변수는 관계가 없음
- 한편, 상관계수는 두 변수의 측정단위 변화에 따라 크기 및 부호가 영향을 받지 않으므로(이를 unit-free라고 함) 많이 이용됨

- (예) 제3장의 기술통계량 계산 시 사용한 자료(describe.xlsx)로 공분산 및 상관계수를 구해보라
- <http://kanggc.ipstime.org/book/data/describe.xlsx>로 파일을 다운로드(왼쪽 그림)
- 데이터-분석-데이터 분석을 실행하면 나타나는 통계 데이터분석 대화상자에서 상관분석(공분산분석)을 실행
- 이때 계열이름이 있는 셀과 자료가 입력되어 있는 셀을 모두 선택한 후 아래 그림과 같은 상관분석 대화상자에서 첫째 행 이름표 사용에 ☒

상관 분석

입력
입력 범위(I):

데이터 방향:
☒ 열(C)
☐ 행(R)

☒ 첫째 행 이름표 사용(L)

출력 옵션
☐ 출력 범위(O):
☒ 새로운 워크시트(P):
☐ 새로운 통합 문서(W)

확인 취소 도움말(H)

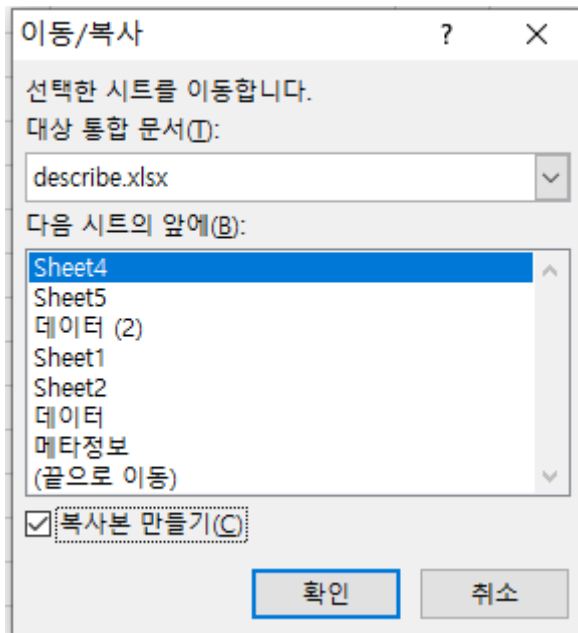
- 모든 공분산이 상관계수와 마찬가지로 양의 부호를 가지고 있고, 대각행렬에 있는 값은 각 변수의 분산

	A	B	C	D	E
1		소비자물가상승률	주택담보대출금리	GDP성장률	GDP디플레이터상승률
2	소비자물가상승률	1.215822222			
3	주택담보대출금리	1.122755556	1.551715556		
4	GDP성장률	0.279866667	0.781333333	2.869066667	
5	GDP디플레이터상승률	0.568488889	0.615755556	0.1352	1.261155556
6					

- 모든 상관계수가 양의 부호를 가지고 있고, 소비자물가상승률과 주택담보대출금리 간 상관계수가 0.817로 가장 크고, GDP성장률과 GDP디플레이터상승률 간 상관계수가 0.071로 가장 작음

	A	B	C	D	E
1		소비자물가상승률	주택담보대출금리	GDP성장률	GDP디플레이터상승률
2	소비자물가상승률	1			
3	주택담보대출금리	0.81741785	1		
4	GDP성장률	0.149846229	0.370305341	1	
5	GDP디플레이터상승률	0.459095054	0.44016748	0.071075847	1

- 상승률, 이자율, 성장률의 측정 단위를 현재 %에서 100으로 나누어 준 후 공분산 및 상관계수를 계산
- 데이터 시트를 마우스 오른쪽쪽을 클릭하여 이동/복사를 선택하면 나타나는 이동/복사 대화상자에서 복사본 만들기에 ☒하고 확인
- F2셀에 식 =B2/100을 입력하여 %를 100으로 나누어 주고, F2셀을 선택하고 I2셀까지 마우스를 끌어서 클릭하여 복사
- F2셀부터 I2셀까지 블록으로 선택하고 I16셀까지 마우스를 끌어서 클릭하여 복사
- F2셀부터 I16셀까지 블록으로 선택하고 복사를 하고 B2셀에서 마우스 오른쪽을 클릭하여 선택하여 붙여넣기-값을 선택하고 확인을 클릭
- 공분산과 상관계수를 계산해 보면 공분산의 크기는 원래 크기의 10000분의 1로 축소되고, 상관계수의 크기는 변화 없음



	A	B	C	D	E
1		소비자물가상승률	주택담보대출금리	GDP성장률	GDP디플레이터상승률
2	소비자물가상승률	0.000121582			
3	주택담보대출금리	0.000112276	0.000155172		
4	GDP성장률	2.79867E-05	7.81333E-05	0.000286907	
5	GDP디플레이터상승률	5.68489E-05	6.15756E-05	0.00001352	0.000126116

	A	B	C	D	E
1		소비자물가상승률	주택담보대출금리	GDP성장률	GDP디플레이터상승률
2	소비자물가상승률	1			
3	주택담보대출금리	0.81741785	1		
4	GDP성장률	0.149846229	0.370305341	1	
5	GDP디플레이터상승률	0.459095054	0.44016748	0.071075847	1

b3-ch5-1-rev.R

```
library(openxlsx)
```

```
df<-read.xlsx("http://kanggc.ipetime.org/book/data/describe-e.xlsx")
```

```
df_dat<-data.matrix(df)
```

```
cpi<-df_dat[,2]
```

```
interest<-df_dat[,3]
```

```
gdp<-df_dat[,4]
```

```
deflator<-df_dat[,5]
```

```
df_new<-cbind(cpi, interest, gdp, deflator)
```

```
df_new
```

```
(var<-var(df_new))
```

```
df_var<-data.matrix(var)
```

```
(sd<-sqrt(diag(df_var)))
```

```
corr<-cor(df_new)
```

```
corr
```

```
df_new_1<-cbind(cpi/100, interest/100, gdp/100, deflator/100)
```

```
(var_new<-var(df_new_1))
```

```
(corr<-cor(df_new_1))
```

```
> (var<-var(df_new))
```

	cpi	interest	gdp	deflator
cpi	1.3026667	1.2029524	0.2998571	0.6090952
interest	1.2029524	1.6625524	0.8371429	0.6597381
gdp	0.2998571	0.8371429	3.0740000	0.1448571
deflator	0.6090952	0.6597381	0.1448571	1.3512381

```
> corr
```

	cpi	interest	gdp	deflator
cpi	1.0000000	0.8174179	0.14984623	0.45909505
interest	0.8174179	1.0000000	0.37030534	0.44016748
gdp	0.1498462	0.3703053	1.00000000	0.07107585
deflator	0.4590951	0.4401675	0.07107585	1.00000000

```
> (var_new<-var(df_new_1))
```

	[,1]	[,2]	[,3]	[,4]
[1,]	1.302667e-04	1.202952e-04	2.998571e-05	6.090952e-05
[2,]	1.202952e-04	1.662552e-04	8.371429e-05	6.597381e-05
[3,]	2.998571e-05	8.371429e-05	3.074000e-04	1.448571e-05
[4,]	6.090952e-05	6.597381e-05	1.448571e-05	1.351238e-04

```
> (corr<-cor(df_new_1))
```

	[,1]	[,2]	[,3]	[,4]
[1,]	1.0000000	0.8174179	0.14984623	0.45909505
[2,]	0.8174179	1.0000000	0.37030534	0.44016748
[3,]	0.1498462	0.3703053	1.00000000	0.07107585
[4,]	0.4590951	0.4401675	0.07107585	1.00000000

I. Excel 단순회귀분석

II. R 단순회귀분석

1. 회귀분석

- 회귀분석이란 한 변수와 다른 변수 사이의 관계를 분석하는 방법
- 회귀분석에서 변수는 영향을 주는 독립변수와 영향을 받는 종속변수로 구분
- 독립변수는 확정변수로 가정하고 종속변수는 확률변수로 가정
- 회귀분석의 목적은 독립변수의 주어진 값으로 종속변수의 평균값을 예측

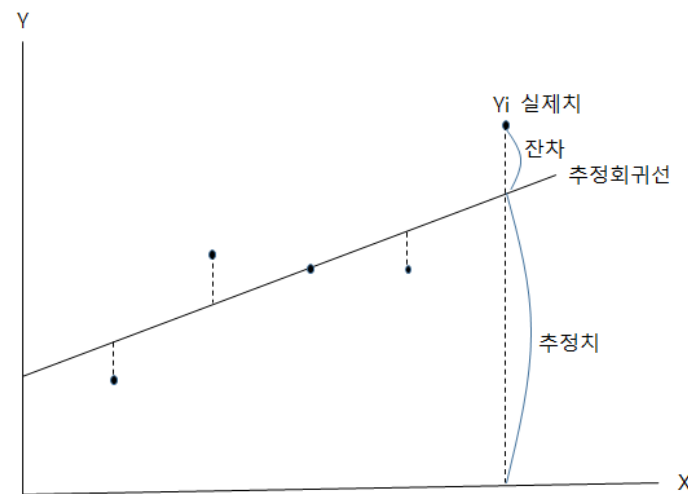
2. 보통최소자승법

- 잔차(=실제치-추정치)의 합계가 최소가 되도록 하는 것이 바람직함
- 잔차의 합이 0이 되는 식은 유일하지가 않음
- 잔차의 제곱의 합이 최소가 되게 하는 회귀식을 구하게 되는데 이러한 추정방법을 보통최소자승법 (Ordinary Least Squares: OLS)이라고 함
- 아래 그림은 단순회귀모형, 추정 회귀선 및 잔차를 나타내 주고 있음

- 모형 : $Y_i = \beta_0 + \beta_1 X_i + u_i$
- 추정된 회귀선 : $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$
- 잔차 : $e_i = Y_i - \hat{Y}_i$

- 단순회귀모형 회귀계수의 추정식은 각각 다음과 같음

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n X_i Y_i - \bar{X} \sum_{i=1}^n Y_i}{\sum_{i=1}^n X_i^2 - \bar{X} \sum_{i=1}^n X_i} = \frac{cov(X, Y)}{var(X)} \quad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$



(1) 추정식을 이용한 계산

- (예 1) X 와 Y에 대한 데이터가 아래 좌측과 같을 때 최소자승법으로 추정한 회귀계수를 각각 계산해 보라
 - A2부터 A6셀에 X 데이터를 입력하고, B2부터 B6셀에 Y 데이터를 입력
 - A7에 =sum(a2:a6)을 입력하여 $\sum_{i=1}^5 X_i$ 를 구하고, A8에 =average(a2:a6)를 입력하여 $\bar{X}$ 를 구함
 - 동일한 방법으로 B7에 $\sum_{i=1}^5 Y_i$ 를 구하고, B8에 $\bar{Y}$ 를 구함
 - C2부터 C6셀에 X^2 를 계산하고, D2부터 D6셀에 XY 를 계산
 - C7에 sum(c2:c6)을 입력하여 $\sum_{i=1}^5 X_i^2$ 를 구하고, D7에 =sum(d2:d6)을 입력하여 $\sum_{i=1}^5 X_i Y_i$ 를 구함
 - $\hat{\beta}_1 = \frac{\sum_{i=1}^5 X_i Y_i - \bar{X} \sum_{i=1}^5 Y_i}{\sum_{i=1}^5 X_i^2 - \bar{X} \sum_{i=1}^5 X_i}$ 를 계산하기 위해 C10에 =(D7-A8*B7)/(C7-A8*A7)을 입력하면 1.4가 계산됨
 - $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$ 를 계산하기 위해 C11에 =B8-C10*A8을 입력하면 0.4가 계산됨

X	Y
2	4
3	4
4	6
5	6
6	10

	A	B	C	D	E	F
1	X	Y	X^2	$X*Y$		
2	2	4	4	8		
3	3	4	9	12		
4	4	6	16	24		
5	5	6	25	30		
6	6	10	36	60		
7	20	30	90	134		
8	4	6				
9	=sum(a2:a6)	=average(a2:a6)	=sum(c2:c6)	=sum(d2:d6)		
10		beta1=	1.4			
11		beta0=	0.4			

(2) 행렬추정식을 이용한 계산

- 행렬을 이용하면 회귀계수는 다음과 같이 추정할 수 있음

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = (X'X)^{-1} X'Y$$

- A14부터 B18셀에 X 행렬을 만드는데, A14부터 A18까지는 상수항 $\hat{\beta}_0$ 추정에 필요한 1을 입력하고, B14부터 B18까지는 $\hat{\beta}_1$ 추정에 필요한 X 데이터를 A2부터 A6셀에 복사해 옴
- E14부터 H15셀에 X의 전치행렬을 만들고, A22부터 B23에 $X'X$, A22부터 B23에 $(X'X)^{-1}$, A26부터 A27에 $X'Y$ 를 만듦
- D26부터 D27을 선택하고 =MMULT(D22:E23,A26:A27) 입력하여 Ctrl + shift + Enter를 누르면 회귀계수를 계산해 줌

	A	B	C	D	E	F	G	H
1	X	Y	X^2	X*Y				
12								
13	X			X'				
14	1	2		1	1	1	1	1
15	1	3		2	3	4	5	6
16	1	4						
17	1	5						
18	1	6						
19								
20								
21	X'X			(X'X)^-1				
22	5	20		1.8	-0.4			
23	20	90		-0.4	0.1			
24								
25	X'Y							
26	30		beta0=	0.4				
27	134		beta1=	1.4				

- 데이터-분석-데이터 분석을 실행하면 나타나는 통계 데이터분석 대화상자에서 회귀분석을 실행하면 아래 좌측 그림과 같은 회귀분석 대화상자가 나타남
- 먼저 종속변수의 데이터 범위를 나타내는 Y축 입력 범위에 \$B\$2:\$B\$6을 입력
- 독립변수의 데이터 범위를 나타내는 X축 입력 범위에 \$A\$2:\$A\$6을 입력
- 회귀 분석의 결과를 인쇄할 영역을 나타내는 출력 범위에 \$B\$29를 입력하고 확인을 누르면 아래 우측 그림과 같은 회귀 분석 결과를 출력

[illegible]

- | | A | B | C |
|----|------|--------------|--------------|
| 1 | 연도 | 국민총소득 | 최종소비지출 |
| 2 | 2000 | 630,614.30 | 413,461.20 |
| 3 | 2001 | 683,447.10 | 460,668.30 |
| 4 | 2002 | 758,862.60 | 515,616.00 |
| 5 | 2003 | 807,778.00 | 535,967.40 |
| 6 | 2004 | 874,238.70 | 562,020.30 |
| 7 | 2005 | 912,608.60 | 602,345.40 |
| 8 | 2006 | 962,446.60 | 643,408.00 |
| 9 | 2007 | 1,040,091.80 | 691,740.40 |
| 10 | 2008 | 1,104,414.30 | 740,804.60 |
| 11 | 2009 | 1,148,981.80 | 769,588.60 |
| 12 | 2010 | 1,266,579.80 | 819,821.20 |
| 13 | 2011 | 1,340,529.80 | 873,522.60 |
| 14 | 2012 | 1,391,595.50 | 911,938.20 |
| 15 | 2013 | 1,439,644.40 | 942,267.20 |
| 16 | 2014 | 1,490,763.90 | 972,924.90 |
| 17 | 2015 | 1,568,383.10 | 1,006,005.60 |
| 18 | 2016 | 1,639,066.50 | 1,047,482.40 |

	A	B	C	D	E	F	G	H	I
1	요약 출력								
2									
3	회귀분석 통계량								
4	다중 상관	0.99871							
5	결정계수	0.997421							
6	조정된 결	0.99725							
7	표준 오차	10600.86							
8	관측수	17							
9									
10	분산 분석								
11		자유도	제곱합	제곱 평균	F 비	유의한 F			
12	회귀	1	6.52E+11	6.52E+11	5802.256	7.81E-21			
13	잔차	15	1.69E+09	1.12E+08					
14	계	16	6.54E+11						
15									
16		계수	표준 오차	t 통계량	P-값	하위 95%	상위 95%	하위 95.0%	상위 95.0%
17	Y 절편	30545.6	9609.735	3.17861	0.006231	10062.94	51028.27	10062.94	51028.27
18	X 1	0.629081	0.008259	76.17254	7.81E-21	0.611478	0.646684	0.611478	0.646684
19									
20		2017년 소비							
21		1099983							

b3-ch5-2.R

```
x<-c(2,3,4,5,6)
y<-c(4,4,6,6,10)
x:y
(n<-length(x))
```

```
(sumx<-sum(x));(sumy<-sum(y))
(mx=mean(x));(my=mean(y))
(xy<-x*y)
(sumxy<-sum(xy));(sumxsq<-sum(x^2));(sumysq<-sum(y^2))
beta1<-(sumxy-mx*sumy)/(sumxsq-mx*sumx)
beta0<-my-beta1*mx
beta0;beta1
(b1hat<-cov(x,y)/var(x))
(b0hat<-mean(y)-b1hat*mean(x))
```

```
X<-matrix(c(1,1,1,1,1,2,3,4,5,6), nrow=5, ncol=2)
Y<-matrix(c(4,4,6,6,10), nrow=5)
N<-length(Y)
X
```

```
(XPX<-t(X)%*%X)
(XPXINV<-solve(XPX))
(XPY<-t(X)%*%Y)
(beta<-XPXINV%*%XPY)
```

```
ols<-lm(y~x)
summary(ols)
```

```
> (sumx<-sum(x))
[1] 20
```

```
> (sumy<-sum(y))
[1] 30
```

```
> (mx=mean(x))
[1] 4
```

```
> (my=mean(y))
[1] 6
```

```
> (sumxy<-sum(xy))
[1] 134
```

```
> (sumxsq<-sum(x^2))
[1] 90
```

```
> (sumysq<-sum(y^2))
[1] 204
```

```
> beta0;beta1
[1] 0.4
[1] 1.4
```

```
> (b1hat<-cov(x,y)/var(x))
[1] 1.4
```

```
> (b0hat<-mean(y)-b1hat*mean(x))
[1] 0.4
```

```
> X
      [,1] [,2]
[1,]  1    2
[2,]  1    3
[3,]  1    4
[4,]  1    5
[5,]  1    6
```

```
> (XPX<-t(X)%*%X)
      [,1] [,2]
[1,]  5    20
[2,]  20   90
```

```
> (XPXINV<-solve(XPX))
      [,1] [,2]
[1,]  1.8 -0.4
[2,] -0.4  0.1
```

```
> (XPY<-t(X)%*%Y)
      [,1]
[1,]  30
[2,] 134
```

```
> (beta<-XPXINV%*%XPY)
      [,1]
[1,]  0.4
[2,]  1.4
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.400	1.625	0.246	0.8214
x	1.400	0.383	3.656	0.0354 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.211 on 3 degrees of freedom
Multiple R-squared: 0.8167, Adjusted R-squared: 0.7556
F-statistic: 13.36 on 1 and 3 DF, p-value: 0.03535

b3-ch5-4.R

```
library(openxlsx)

df<-read.xlsx("http://kanggc.iptime.org/book/data/consumption-e.xlsx")

year<-df[,1]
gdp<-df[,2]
consumption<-df[,3]

y<-ts(gdp, start=c(2000), frequency=1)
c<-ts(consumption, start=c(2000), frequency=1)

n<-length(y)

ny<-y[2:n]
nc<-c[2:n]
lagc<-c[1:n-1]

lm_s<-lm(c~y)
ols_s<-summary(lm_s)
ols_s
y0<-1700000
(chat_s<-summary(lm_s)$coef[1]+summary(lm_s)$coef[2]*y0)
*y0)
다음 장에 계속
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.055e+04	9.610e+03	3.179	0.00623 **
y	6.291e-01	8.259e-03	76.173	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10600 on 15 degrees of freedom
Multiple R-squared: 0.9974, Adjusted R-squared: 0.9972
F-statistic: 5802 on 1 and 15 DF, p-value: < 2.2e-16

```
> y0<-1700000
```

```
> (chat_s<-summary(lm_s)$coef[1]+summary(lm_s)$coef[2]*y0)
[1] 1099983
```

I. Excel 다중회귀분석

II. R 다중회귀분석

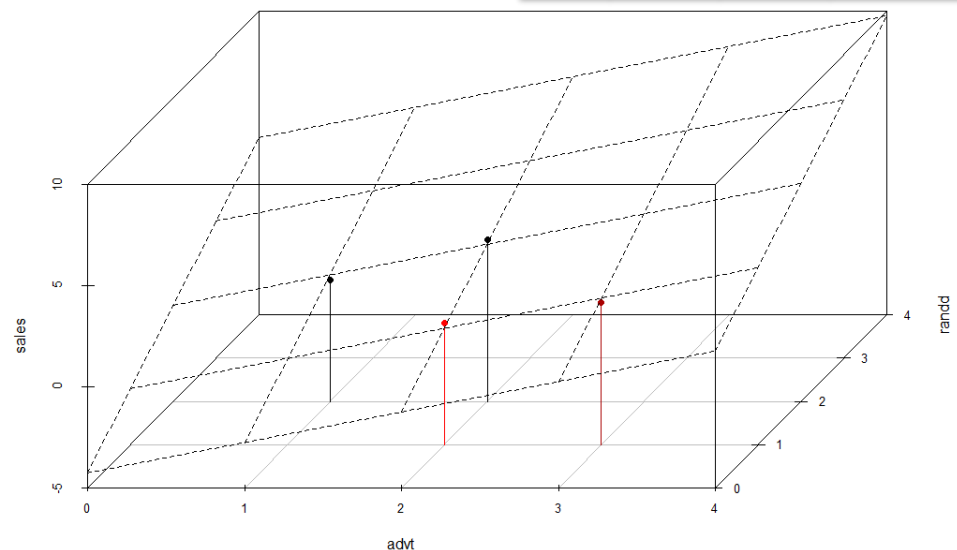
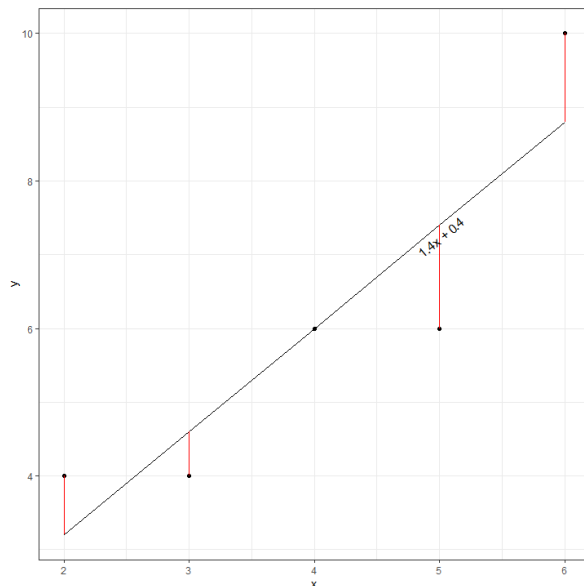
1. 다중회귀모형

- 다중회귀모형이란 설명변수가 2개 이상인 회귀모형을 말함
- 독립변수 또는 종속변수의 시차변수가 설명변수로 포함된 동태모형일 경우 회귀분석에서 다음 사항에 유의해야 함
 - 시차변수 데이터는 기존의 독립변수 또는 종속변수의 데이터를 이용
 - 회귀모형의 추정 시 사용되는 독립변수 및 종속변수의 관측치 개수를 동일하게 해야 함

2. 보통최소자승법

- 좌측 그림은 단순회귀모형 및 보통최소자승법을 시각적으로 보여 주고 있음
 - 단순회귀모형 : $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_t$
- 우측 그림은 기업의 매출액(sales)은 홍보비(advt) 및 연구개발비(randd)의 함수라는 다중회귀모형 및 보통최소자승법을 시각적으로 보여 주고 있음
 - 다중회귀모형 : $sales_i = \beta_0 + \beta_1 advt_i + \beta_2 randd_i + \varepsilon_t$

기업	매출액	홍보비	연구개발비
1	1	1	2
2	1	2	1
3	2	3	1
4	3	2	2



- (예) 소비(단위 : 10억 원)와 소득(단위 : 10억 원)의 자료(consumption.xlsx)로 현재소비(C_t)는 현재소득(Y_t)과 전년도 소비(C_{t-1})의 함수라는 다중회귀모형을 추정하고, 2017년도 소득이 1700조라고 할 때 2017년도 소비를 예측해 보라

$$C_t = \beta_0 + \beta_1 Y_t + \beta_2 C_{t-1} + \varepsilon_t$$

- <http://kanggc.ipstime.org/book/data/consumption.xlsx>로 파일을 다운로드
- 위의 다중회귀모형은 현재소비(C_t)의 시차변수인 전년도 소비(C_{t-1})가 독립변수에 포함되어 있으므로 이 데이터를 만들어야 함
- E2셀부터 E18셀까지 현재소득(Y_t)을 복사하고, 전년도 소비(C_{t-1})를 만들기 위해 F3셀에 식 =C2를 입력하고, 이를 F18셀까지 복사
- 위 모형은 동태모형이므로 회귀분석에서 2001년부터 2016년까지 데이터가 이용되므로 데이터-분석-데이터 분석을 실행하여 통계 데이터분석 대화상자에서 회귀분석을 선택하면 나타나는 회귀분석 대화상자에서 Y축 입력범위에 \$C\$3:\$C\$18을 선택하고, X축 입력 범위에 \$E\$3:\$F\$18을 선택하여 확인을 클릭하면 다음 그림과 같은 소비함수 추정 결과가 나타남
- 2017년 소비를 예측하기 위하여 B22셀에
식 =B17+B18*1700000+B19*1047482를
입력하면 2017년 소비 예측치가 1092084임

	A	B	C	D	E	F	G	H	I
1	요약 출력								
2									
3	회귀분석 통계량								
4	다중 상관	0.999157							
5	결정계수	0.998315							
6	조정된 결	0.998056							
7	표준 오차	8390.698							
8	관측수	16							
9									
10	분산 분석								
11		자유도	제곱합	제곱 평균	F 비	유의한 F			
12	회귀	2	5.42E+11	2.71E+11	3851.938	9.38E-19			
13	잔차	13	9.15E+08	70403808					
14	계	15	5.43E+11						
15									
16		계수	표준 오차	t 통계량	P-값	하위 95%	상위 95%	하위 95.0%	상위 95.0%
17	Y 절편	40295.23	8523.49	4.727551	0.000395	21881.35	58709.11	21881.35	58709.11
18	X 1	0.33455	0.105585	3.168532	0.007403	0.106447	0.562653	0.106447	0.562653
19	X 2	0.461157	0.167887	2.74683	0.016638	0.098459	0.823855	0.098459	0.823855
20									
21			2017년 소비						
22			1092084						

b3-ch5-4.R

```
library(openxlsx)
df<-read.xlsx("http://kanggc.ipstime.org/book/data/consumption-e.xlsx")
year<-df[,1]
gdp<-df[,2]
consumption<-df[,3]
y<-ts(gdp, start=c(2000), frequency=1)
c<-ts(consumption, start=c(2000), frequency=1)
n<-length(y)
(ny<-y[2:n])
(nc<-c[2:n])
(lagc<-c[1:n-1])
(data<-cbind(ny,nc,lagc))
lm_s<-lm(c~y)
ols_s<-summary(lm_s)
ols_s
y0<-1700000
(chat_s<-summary(lm_s)$coef[1]+summary(lm_s)$coef[2]*y0)
lm_m<-lm(nc~ny+lagc)
ols_m<-summary(lm_m)
ols_m
(chat_m<-summary(lm_m)$coef[1]+summary(lm_m)$coef[2]*y0+summary(lm_m)$coef[3]*c[17])
```

```
> (data<-cbind(ny,nc,lagc))
```

	ny	nc	lagc
[1,]	683447.1	460668.3	413461.2
[2,]	758862.6	515616.0	460668.3
[3,]	807778.0	535967.4	515616.0
[4,]	874238.7	562020.3	535967.4
[5,]	912608.6	602345.4	562020.3
[6,]	962446.6	643408.0	602345.4
[7,]	1040091.8	691740.4	643408.0
[8,]	1104414.3	740804.6	691740.4
[9,]	1148981.8	769588.6	740804.6
[10,]	1266579.8	819821.2	769588.6
[11,]	1340529.8	873522.6	819821.2
[12,]	1391595.5	911938.2	873522.6
[13,]	1439644.4	942267.2	911938.2
[14,]	1490763.9	972924.9	942267.2
[15,]	1568383.1	1006005.6	972924.9
[16,]	1639066.5	1047482.4	1006005.6

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	4.030e+04	8.523e+03	4.728	0.000395 ***
ny	3.346e-01	1.056e-01	3.169	0.007403 **
lagc	4.612e-01	1.679e-01	2.747	0.016638 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8391 on 13 degrees of freedom
Multiple R-squared: 0.9983, Adjusted R-squared: 0.9981
F-statistic: 3852 on 2 and 13 DF, p-value: < 2.2e-16

```
> (chat_m<-summary(lm_m)$coef[1]+summary(lm_m)$coef[2]*y0+summary(lm_m)$coef[3]*c[17])
[1] 1092084
```