

Coefficient Investigation

David Leslie

Sunday, April 12, 2015

- This code assumes an array filled with spatially autocorrelated data has already been generated.

Now that the data has been generated, let's see how the accuracy of the model coefficients changes as the spatial range increases (Remembering that they should be close to the expected values of: $x_1=2$, $x_2=1$, and $x_3=3$).

```
# Load data
load("~/R/EVSS695/spatial_cv/resultsVarStagScaleStep.Rdata")

# Must match numMat from generated data
numMat = 100

# Must match numReps from generated data
numReps = 10

# ust match spat_range from generated data
spat_range = seq(0.001, 60, length.out=numMat)

# ust match var_range from generated data
var_range = 1:10

# Set numbers of columns in coefficients array
numColsCoef = 5

# Create array filled with coefficient values
coefVals = array(NA, dim=c(numMat, numColsCoef, numReps), dimnames = list(1:numMat, c('x1_dif', 'x2_dif', 'x3_dif', 'spat_range', 'var_range'), 1:numReps))

# Fill array with coefficient values
for(i in 1:numReps) {
  for(j in 1:numMat) {
    model = glm(y ~ x1 + x2 + x3, data = as.data.frame(resultsArray[,j,i]))
    coefVals[j, 1, i] = 2 - coef(model)[2]
    coefVals[j, 2, i] = 1 - coef(model)[3]
    coefVals[j, 3, i] = 3 - coef(model)[4]
    coefVals[j, 4, i] = spat_range[j]
    coefVals[j, 5, i] = var_range[i]
  }
}

# Check array
head(coefVals[, , 1])
```

##	x1_dif	x2_dif	x3_dif	spat_range	var_range
## 1	-0.000364776	0.007487960	0.001283645	0.0010000	1
## 2	-0.001672900	-0.007414070	0.001695848	0.6070505	1
## 3	-0.003679532	-0.025158432	-0.004636865	1.2131010	1
## 4	-0.012699380	-0.012731045	0.018647603	1.8191515	1

```
## 5 -0.001277530 -0.017973304 -0.005635941 2.4252020 1
## 6 0.024120265 -0.003303756 -0.043411985 3.0312525 1
```

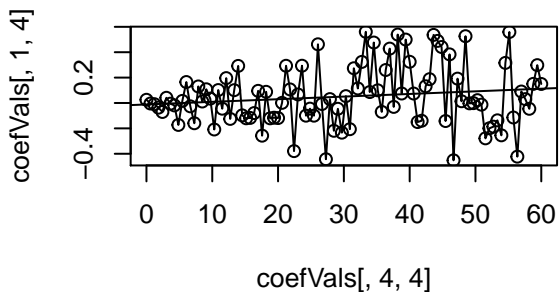
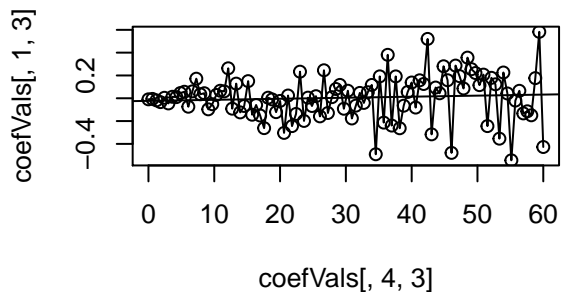
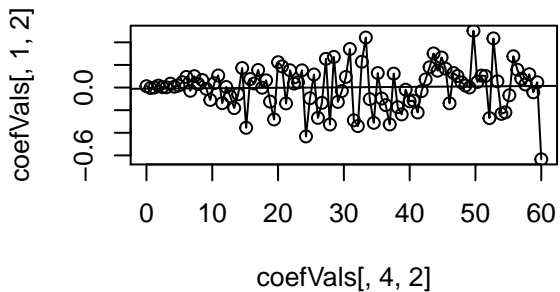
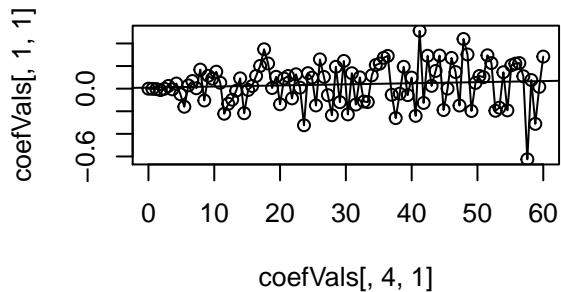
```
head(coefVals[, ,5])
```

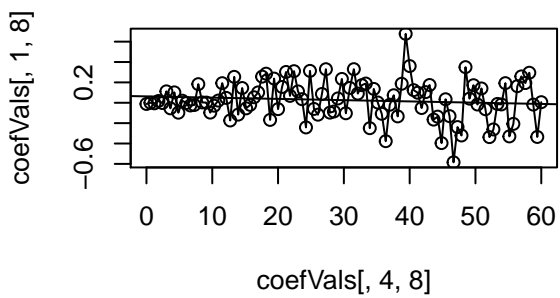
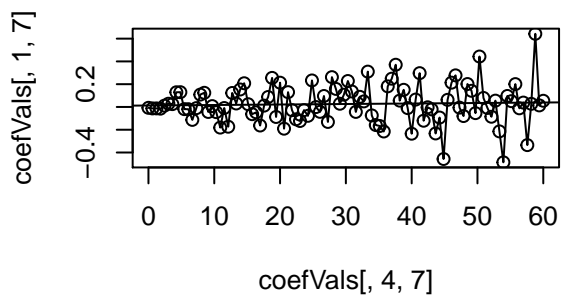
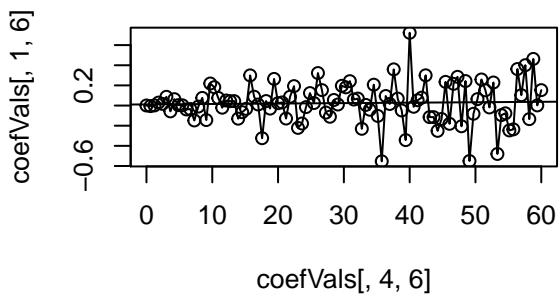
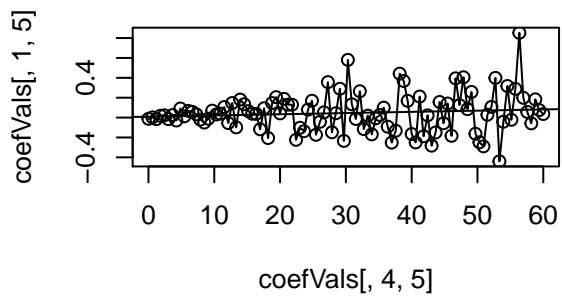
```
##          x1_dif          x2_dif          x3_dif spat_range var_range
## 1 -0.012276188 -0.0192985607 -0.0040470746 0.0010000 5
## 2 0.002150259 -0.0002968889 0.0051174141 0.6070505 5
## 3 -0.015388971 0.0046331308 0.0002646123 1.2131010 5
## 4 0.017822810 0.0163953828 -0.0152363895 1.8191515 5
## 5 0.022723972 -0.0253002927 -0.0042946876 2.4252020 5
## 6 -0.015619837 -0.0156150866 -0.0712717121 3.0312525 5
```

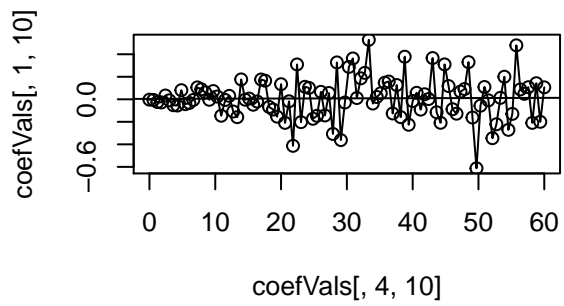
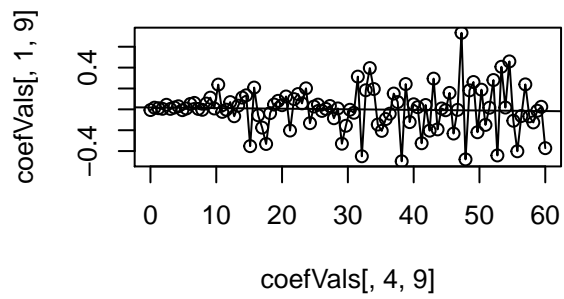
```
head(coefVals[, ,10])
```

```
##          x1_dif          x2_dif          x3_dif spat_range var_range
## 1 -0.003048424 0.0049080826 -0.003078637 0.0010000 10
## 2 -0.005351462 0.0005459124 0.001954698 0.6070505 10
## 3 -0.024410876 0.0129280014 -0.006218236 1.2131010 10
## 4 -0.027896899 -0.0446239591 -0.008066931 1.8191515 10
## 5 0.033434246 -0.0297341170 0.004654499 2.4252020 10
## 6 -0.009138082 -0.0737190095 -0.017273967 3.0312525 10
```

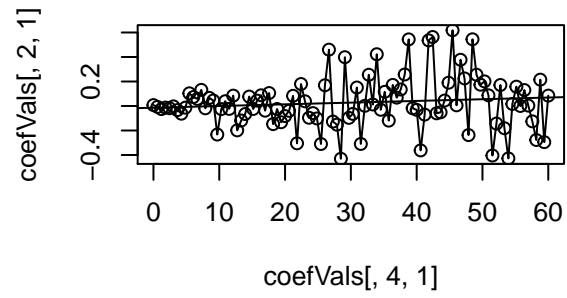
X1 Coefficient Difference Plots



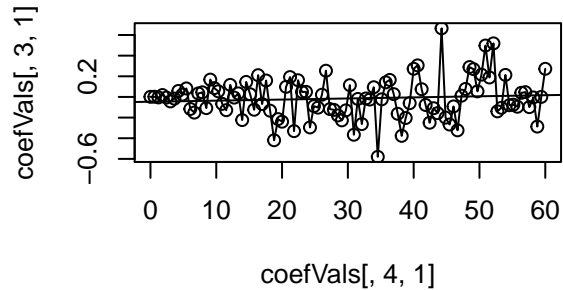




X2 Coefficient Difference Plots



X3 Coefficient Difference Plots



From the plots, it appears that the model coefficients are somewhat accurate. Now let's look at how the precision changes.

```
"
coefVals = matrix(NA, nrow=numMat, ncol=3, dimnames=list(1:numMat, c('x1_dif', 'x2_dif', 'x3_dif')))
for(i in 1:numMat) {
  model = glm(y ~ x1 + x2 + x3, data = as.data.frame(resultsArray[,i]))
  coefVals[i,] = coef(model)[2:4]
}

coefVals[,1] = abs(2 - coefVals[,1])
coefVals[,2] = abs(1 - coefVals[,2])
coefVals[,3] = abs(3 - coefVals[,3])
coefVals = cbind(coefVals, spat_range)

# Plot precision
par(mfrow=c(2,2))
plot(coefVals[,4], coefVals[,1], type='o')
abline(lm(coefVals[,1] ~ coefVals[,4]))

plot(coefVals[,4], coefVals[,2], type='o')
abline(lm(coefVals[,2] ~ coefVals[,4]))
```

```
plot(coefVals[,4], coefVals[,3], type='o')
abline(lm(coefVals[,3] ~ coefVals[,4]))
"
```

```
## [1] "\ncoefVals = matrix(NA, nrow=numMat, ncol=3, dimnames=list(1:numMat,c('x1_dif', 'x2_dif', 'x3_dif')))"
```