# Make data matrix
X = matrix(rnorm(3000),100,30)
# Set noise level
sigma = 4
# Define model
Y = 3 - 5*X[,1] + X[,2] - 4*X[,3] + 2*X[,4] + sigma*rnorm(100)

# Consider nested model sequence
full = lm(Y ~ ., data = as.data.frame(X))
null = lm(Y ~ 1, data = as.data.frame(X))
s = sqrt(sum(full$residuals^2)/full$df.residual)

# Setting k = 0 forces to always add variables
step(null,scope=list(lower=null,upper=full),direction="forward", scale = s^2,
data = as.data.frame(X),k=0)

# We now have the order with which the variables which enter in the
# model which key in below. If you run the script, you will need to
# input the ordering you got.
variables.to.enter = c(1, 3, 4, 2, 5, 18, 10, 7, 20, 12, 17, 21, 14, 24, 11, 8, 25, 13, 27, 19, 30, 29, 6, 22, 16, 23, 28, 9, 15, 26)

# Check cv error
rr = selection_criteria(Y, cbind(1,X[,variables.to.enter]))
cv.error = rr$cv
cv.error
plot(cv.error)

# Check Cp
Cp.stat = rr$Cp
Cp.stat
plot(Cp.stat)
Figure 1: CV error as a function of model size
Figure 2: $C_p$ statistic as a function of model size