

Gaussian_Process_Code

Chiwan Kim

2/3/2020

##Part 1: Standard Gaussian Process

1-1: Fitting

```
library(rstan)
```

```
## Loading required package: StanHeaders
```

```
## Loading required package: ggplot2
```

```
## rstan (Version 2.19.3, GitRev: 2e1f913d3ca3)
```

```
## For execution on a local, multicore CPU with excess RAM we recommend calling
```

```
## options(mc.cores = parallel::detectCores()).
```

```
## To avoid recompilation of unchanged Stan programs, we recommend calling
```

```
## rstan_options(auto_write = TRUE)
```

```
source("gp.utility.R")
```

```
# Fitting GP model
```

```
stan_dat <- read_rdump('Financial_Data_Put_American.R')
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
##      filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      intersect, setdiff, setequal, union
```

```
## Parsed with column specification:
```

```
## cols(
```

```
##   .default = col_double(),
```

```
##   date = col_character(),
```

```
##   symbol = col_character(),
```

```
##   exdate = col_character(),
```

```
##   cp_flag = col_character(),
```

```
##   ticker = col_character(),
```

```
##   exercise_style = col_character()
```

```
## )
```

```
## See spec(...) for full column specifications.
```

```
## Warning: 98350 parsing failures.
```

```
##      row      col expected actual
```

```

## 142894 6/21/2019 a double FALSE '~/projects/Independent_Study/spy_spx_(2019.06.01~2019.06.30)_Puts
## 142894 9/20/2019 a double FALSE '~/projects/Independent_Study/spy_spx_(2019.06.01~2019.06.30)_Puts
## 142894 12/20/2019 a double FALSE '~/projects/Independent_Study/spy_spx_(2019.06.01~2019.06.30)_Puts
## 142895 6/21/2019 a double FALSE '~/projects/Independent_Study/spy_spx_(2019.06.01~2019.06.30)_Puts
## 142895 9/20/2019 a double FALSE '~/projects/Independent_Study/spy_spx_(2019.06.01~2019.06.30)_Puts
## .....
## See problems(...) for more details.

## Loading required package: limSolve

##
## Attaching package: 'limSolve'

## The following object is masked from 'package:ggplot2':
##
## resolution

## Loading required package: futile.logger

## Welcome to ragtop. Logging can be enabled with commands such as
## futile.logger::flog.threshold(futile.logger::INFO, name='ragtop.calibration')

## Registered S3 method overwritten by 'quantmod':
## method from
## as.zoo.data.frame zoo

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

```


[illegible]

[illegible]


```

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

fit_gp_SGP_American <- stan(file="gp-fit-6dimension_withBS.stan", data=stan_dat,
                           iter=100, chains=1);

## DIAGNOSTIC(S) FROM PARSER:
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.

## Trying to compile a simple C file

## Running /usr/lib64/R/bin/R CMD SHLIB foo.c
## gcc -m64 -I"/usr/include/R" -DNDEBUG -I"/usr/lib64/R/library/Rcpp/include/" -I"/usr/lib64/R/libra
## In file included from /usr/lib64/R/library/RcppEigen/include/Eigen/Dense:1,
##                  from /usr/lib64/R/library/StanHeaders/include/stan/math/prim/mat/fun/Eigen.hpp:13,
##                  from <command-line>:
## /usr/lib64/R/library/RcppEigen/include/Eigen/Core:82:12: fatal error: new: No such file or directory
##   #include <new>
##           ~~~~~
## compilation terminated.
## make: *** [/usr/lib64/R/etc/Makeconf:167: foo.o] Error 1
##
## SAMPLING FOR MODEL 'gp-fit-6dimension_withBS' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 0.020266 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 202.66 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:

```

```

## Chain 1:
## Chain 1: WARNING: There aren't enough warmup iterations to fit the
## Chain 1:           three stages of adaptation as currently configured.
## Chain 1:           Reducing each adaptation stage to 15%/75%/10% of
## Chain 1:           the given number of warmup iterations:
## Chain 1:           init_buffer = 7
## Chain 1:           adapt_window = 38
## Chain 1:           term_buffer = 5
## Chain 1:
## Chain 1: Iteration:  1 / 100 [ 1%] (Warmup)
## Chain 1: Iteration: 10 / 100 [10%] (Warmup)
## Chain 1: Iteration: 20 / 100 [20%] (Warmup)
## Chain 1: Iteration: 30 / 100 [30%] (Warmup)
## Chain 1: Iteration: 40 / 100 [40%] (Warmup)
## Chain 1: Iteration: 50 / 100 [50%] (Warmup)
## Chain 1: Iteration: 51 / 100 [51%] (Sampling)
## Chain 1: Iteration: 60 / 100 [60%] (Sampling)
## Chain 1: Iteration: 70 / 100 [70%] (Sampling)
## Chain 1: Iteration: 80 / 100 [80%] (Sampling)
## Chain 1: Iteration: 90 / 100 [90%] (Sampling)
## Chain 1: Iteration: 100 / 100 [100%] (Sampling)
## Chain 1:
## Chain 1: Elapsed Time: 34.1436 seconds (Warm-up)
## Chain 1:           44.7943 seconds (Sampling)
## Chain 1:           78.9379 seconds (Total)
## Chain 1:

## Warning: The largest R-hat is 1.25, indicating chains have not mixed.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#r-hat

## Warning: Bulk Effective Samples Size (ESS) is too low, indicating posterior means and medians may be
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#bulk-ess

## Warning: Tail Effective Samples Size (ESS) is too low, indicating posterior variances and tail quant.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#tail-ess

print(fit_gp_SGP_American, pars = c('theta', 'sigma2', 'gamma2'))

## Inference for Stan model: gp-fit-6dimension_withBS.
## 1 chains, each with iter=100; warmup=50; thin=1;
## post-warmup draws per chain=50, total post-warmup draws=50.
##
##           mean    se_mean      sd    2.5%    25%    50%    75%
## theta[1] 50104.27   8844.34  35877.26   40.75 25192.06 40416.64 68015.44
## theta[2]  227.02     9.76   89.98  102.63  153.98  210.65  279.24
## theta[3]   17.31     0.63   4.15   10.24   14.60   16.60   20.36
## theta[4]    0.02     0.00   0.01    0.01    0.02    0.02    0.02
## theta[5] 405784.93 260969.37 801752.67    0.11   43.23  4141.26 534703.72
## theta[6]   11.26     1.67  10.00    1.24    5.08    8.20   14.58
## sigma2     0.00     0.00   0.00    0.00    0.00    0.00    0.00
## gamma2  12287.88  1169.72  8136.02 4030.35  7739.18  9749.47 15591.34
##
##           97.5% n_eff Rhat
## theta[1] 129521.00    16 1.19

```



```

## theta[2]      407.86      85 1.01
## theta[3]      23.84      43 0.98
## theta[4]       0.04      34 1.01
## theta[5] 2198636.27       9 1.08
## theta[6]      35.63      36 0.98
## sigma2        0.00      50 1.05
## gamma2     33164.20      48 0.98
##
## Samples were drawn using NUTS(diag_e) at Thu Mar 26 16:10:07 2020.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).

sum_gp_SGP_American <- extract(fit_gp_SGP_American,permuted=FALSE)

# Predicting from GP model
post_mean_theta_1_SGP <- mean(sum_gp_SGP_American[,1,1]) #theta
post_mean_theta_2_SGP <- mean(sum_gp_SGP_American[,1,2]) #theta
post_mean_theta_3_SGP <- mean(sum_gp_SGP_American[,1,3]) #theta
post_mean_theta_4_SGP <- mean(sum_gp_SGP_American[,1,4]) #theta
post_mean_theta_5_SGP <- mean(sum_gp_SGP_American[,1,5]) #theta
post_mean_theta_6_SGP <- mean(sum_gp_SGP_American[,1,6]) #theta
post_mean_sigma2_SGP <- mean(sum_gp_SGP_American[,1,7]) #sigma2
post_mean_gamma2_SGP <- mean(sum_gp_SGP_American[,1,8]) #gamma2
post_mean_mu_SGP <- stan_dat$blackscholes

# x2 <- as.numeric(unlist(spx_spy_2019_06_30_put_2017_06_500rows_test['strike_price']))
# x2<- cbind(spy_2013_01_01_2013_01_31_put$strike_price[201:300],spy_2013_01_01_2013_01_31_put$impl_vol)
# x2 <- seq(from=-2,to=2,by=0.01)

# x2 <- cbind(seq(from=0,to=1,by=0.01),seq(from=0,to=1,by=0.01))

# test_start <- 323 #06/10
# test_end <- 559 #06/14

test_start <- 560 #06/17
test_end <- 852 #06/20

x.grid_1 <- as.numeric(stan_dat$total_puts_American$forward_price[test_start:test_end])
x.grid_2 <- as.numeric(stan_dat$total_puts_American$strike_price[test_start:test_end])
x.grid_3 <- as.numeric(stan_dat$total_puts_American$impl_volatility[test_start:test_end])
x.grid_4 <- as.numeric(stan_dat$total_puts_American$time_to_exp[test_start:test_end])
x.grid_5 <- as.numeric(stan_dat$total_puts_American$dividend[test_start:test_end])
x.grid_6 <- as.numeric(stan_dat$total_puts_American$interest_rate[test_start:test_end])
x2 <- cbind(x.grid_1,x.grid_2,x.grid_3,x.grid_4,x.grid_5,x.grid_6)

library('qrmtools')
library('ragtop')
blackscholes_2 <- rep(NA,length(x2[,1]))
for (row in 1:nrow(data.frame(x2))) {
  blackscholes_2[row] <- as.numeric(blackscholes(-1,S0=x.grid_1[row],K=x.grid_2[row],r=x.grid_6[row],t=
# blackscholes_2[row] <- Black_Scholes(0,x.grid_1[row],x.grid_6[row],x.grid_3[row],x.grid_2[row],x.gr
})

```

1-2: Predictions

```
# X.grid <- expand.grid(x1 = x.grid_1, x2 = x.grid_2)
```

```
post_data_SGP_American <- list(theta=c(post_mean_theta_1_SGP,post_mean_theta_2_SGP,post_mean_theta_3_SGP),  
# post_data
```

```
pred_gp_SGP <- stan(file="Predictive GP_6dimension_withBS.stan", data=post_data_SGP_American,iter=200, v
```

```
## DIAGNOSTIC(S) FROM PARSER:
```

```
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
```

```
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
```

```
## Trying to compile a simple C file
```

```
## Running /usr/lib64/R/bin/R CMD SHLIB foo.c
```

```
## gcc -m64 -I"/usr/include/R" -DNDEBUG -I"/usr/lib64/R/library/Rcpp/include/" -I"/usr/lib64/R/libra
```

```
## In file included from /usr/lib64/R/library/RcppEigen/include/Eigen/Dense:1,
```

```
## from /usr/lib64/R/library/RcppEigen/include/StanHeaders/include/stan/math/prim/mat/fun/Eigen.hpp:13,
```

```
## from <command-line>:
```

```
## /usr/lib64/R/library/RcppEigen/include/Eigen/Core:82:12: fatal error: new: No such file or directory
```

```
## #include <new>
```

```
## ~~~~~
```

```
## compilation terminated.
```

```
## make: *** [/usr/lib64/R/etc/Makeconf:167: foo.o] Error 1
```

```
##
```

```
## SAMPLING FOR MODEL 'Predictive GP_6dimension_withBS' NOW (CHAIN 1).
```

```
## Chain 1: Iteration: 1 / 200 [ 0%] (Sampling)
```

```
## Chain 1: Iteration: 100 / 200 [ 50%] (Sampling)
```

```
## Chain 1: Iteration: 200 / 200 [100%] (Sampling)
```

```
## Chain 1:
```

```
## Chain 1: Elapsed Time: 0 seconds (Warm-up)
```

```
## Chain 1: 9.49489 seconds (Sampling)
```

```
## Chain 1: 9.49489 seconds (Total)
```

```
## Chain 1:
```

```
##Part2: Bdrycov Gaussian Process
```

2-1: Fitting

```
# Fitting GP model for Bdrycov
```

```
stan_dat <- read_rdump('Financial_Data_Put_American.R')
```

```
## Parsed with column specification:
```

```
## cols(
```

```
## .default = col_double(),
```

```
## date = col_character(),
```

```
## symbol = col_character(),
```

```
## exdate = col_character(),
```

```
## cp_flag = col_character(),
```

```
## ticker = col_character(),
```

```
## exercise_style = col_character()
```

```
## )
```

```
## See spec(...) for full column specifications.
```

```
## Warning: 98350 parsing failures.
```

```
## row col expected actual
```


[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```

## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

fit_gp_Bdrycov_American <- stan(file="gp-fit-6dimension_withBS_Bdrycov.stan", data=stan_dat,
                               iter=100, chains=1);

## Trying to compile a simple C file

## Running /usr/lib64/R/bin/R CMD SHLIB foo.c
## gcc -m64 -I"/usr/include/R" -DNDEBUG -I"/usr/lib64/R/library/Rcpp/include/" -I"/usr/lib64/R/libra
## In file included from /usr/lib64/R/library/RcppEigen/include/Eigen/Dense:1,
##                  from /usr/lib64/R/library/StanHeaders/include/stan/math/prim/mat/fun/Eigen.hpp:13,
##                  from <command-line>:
## /usr/lib64/R/library/RcppEigen/include/Eigen/Core:82:12: fatal error: new: No such file or directory
##   #include <new>
##           ~~~~~
## compilation terminated.
## make: *** [/usr/lib64/R/etc/Makeconf:167: foo.o] Error 1
##
## SAMPLING FOR MODEL 'gp-fit-6dimension_withBS_Bdrycov' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 0.126055 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 1260.55 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:
## Chain 1:
## Chain 1: WARNING: There aren't enough warmup iterations to fit the
## Chain 1:           three stages of adaptation as currently configured.
## Chain 1:           Reducing each adaptation stage to 15%/75%/10% of
## Chain 1:           the given number of warmup iterations:
## Chain 1:           init_buffer = 7
## Chain 1:           adapt_window = 38
## Chain 1:           term_buffer = 5
## Chain 1:
## Chain 1: Iteration:  1 / 100 [  1%] (Warmup)
## Chain 1: Iteration: 10 / 100 [ 10%] (Warmup)
## Chain 1: Iteration: 20 / 100 [ 20%] (Warmup)
## Chain 1: Iteration: 30 / 100 [ 30%] (Warmup)
## Chain 1: Iteration: 40 / 100 [ 40%] (Warmup)
## Chain 1: Iteration: 50 / 100 [ 50%] (Warmup)
## Chain 1: Iteration: 51 / 100 [ 51%] (Sampling)

```



```

## Chain 1: Iteration: 60 / 100 [ 60%] (Sampling)
## Chain 1: Iteration: 70 / 100 [ 70%] (Sampling)
## Chain 1: Iteration: 80 / 100 [ 80%] (Sampling)
## Chain 1: Iteration: 90 / 100 [ 90%] (Sampling)
## Chain 1: Iteration: 100 / 100 [100%] (Sampling)
## Chain 1:
## Chain 1: Elapsed Time: 722.367 seconds (Warm-up)
## Chain 1: 769.165 seconds (Sampling)
## Chain 1: 1491.53 seconds (Total)
## Chain 1:

## Warning: The largest R-hat is 1.18, indicating chains have not mixed.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#r-hat

## Warning: Bulk Effective Samples Size (ESS) is too low, indicating posterior means and medians may be
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#bulk-ess

## Warning: Tail Effective Samples Size (ESS) is too low, indicating posterior variances and tail quant.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#tail-ess

print(fit_gp_Bdrycov_American, pars = c('theta','sigma2','gamma2'))

## Inference for Stan model: gp-fit-6dimension_withBS_Bdrycov.
## 1 chains, each with iter=100; warmup=50; thin=1;
## post-warmup draws per chain=50, total post-warmup draws=50.
##
##               mean      se_mean      sd 2.5%  25%      50%
## theta[1] 3.248405e+05 2.196982e+05 1.452299e+06 0.11  1.50     13.22
## theta[2] 1.010569e+17 9.679251e+16 5.321590e+17 0.07  7.14 1470898.35
## theta[3] 8.184325e+08 6.289533e+08 4.406763e+09 0.13  2.87     172.33
## theta[4] 1.435375e+14 1.155000e+14 8.589929e+14 0.08  7.52     305.79
## theta[5] 1.248538e+15 1.133214e+15 5.547606e+15 0.09  2.02      67.24
## theta[6] 1.675240e+18 1.603950e+18 1.184573e+19 0.21 19.96    10703.10
## sigma2    5.845000e+01 7.940000e+00 5.033000e+01 0.03  3.10      88.29
## gamma2    4.751000e+01 7.890000e+00 4.868000e+01 0.06  0.91      15.41
##
##               75%      97.5% n_eff Rhat
## theta[1] 8.697690e+03 2.128426e+06   44 1.03
## theta[2] 1.793884e+10 1.107864e+18   30 1.01
## theta[3] 3.983270e+03 5.353208e+09   49 0.99
## theta[4] 2.277726e+05 8.808142e+14   55 0.99
## theta[5] 1.243348e+05 1.516744e+16   24 1.03
## theta[6] 3.213058e+08 3.819663e+13   55 1.00
## sigma2    1.025800e+02 1.204000e+02   40 0.99
## gamma2    9.717000e+01 1.175300e+02   38 0.98
##
## Samples were drawn using NUTS(diag_e) at Thu Mar 26 16:39:06 2020.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).

sum_gp_Bdrycov_American <- extract(fit_gp_Bdrycov_American,permuted=FALSE)
# saveRDS(fit_gp,file ="fit_gp_vol50_within50spot_7to19days")

```

```

# Predicting from GP model - 2 dimensional case
post_mean_theta_1_Bdrycov <- mean(sum_gp_Bdrycov_American[,1,1]) #theta
post_mean_theta_2_Bdrycov <- mean(sum_gp_Bdrycov_American[,1,2]) #theta
post_mean_theta_3_Bdrycov <- mean(sum_gp_Bdrycov_American[,1,3]) #theta
post_mean_theta_4_Bdrycov <- mean(sum_gp_Bdrycov_American[,1,4]) #theta
post_mean_theta_5_Bdrycov <- mean(sum_gp_Bdrycov_American[,1,5]) #theta
post_mean_theta_6_Bdrycov <- mean(sum_gp_Bdrycov_American[,1,6]) #theta
post_mean_sigma2_Bdrycov <- mean(sum_gp_Bdrycov_American[,1,7]) #sigma2
post_mean_gamma2_Bdrycov <- mean(sum_gp_Bdrycov_American[,1,8]) #gamma2
post_mean_mu_Bdrycov <- stan_dat$blackscholes

# x2 <- as.numeric(unlist(spx_spy_2019_06_30_put_2017_06_500rows_test['strike_price']))
# x2<- cbind(spy_2013_01_01_2013_01_31_put$strike_price[201:300],spy_2013_01_01_2013_01_31_put$impl_vol)
# x2 <- seq(from=-2,to=2,by=0.01)

# x2 <- cbind(seq(from=0,to=1,by=0.01),seq(from=0,to=1,by=0.01))

```

2-2: Predictions

```

# X.grid <- expand.grid(x1 = x.grid_1, x2 = x.grid_2)

post_data_Bdrycov_American <- list(theta=c(post_mean_theta_1_Bdrycov,post_mean_theta_2_Bdrycov,post_mean_theta_3_Bdrycov,post_mean_theta_4_Bdrycov,post_mean_theta_5_Bdrycov,post_mean_theta_6_Bdrycov),
# post_data

pred_gp_Bdrycov <- stan(file="Predictive_GP_6dimension_withBS_Bdrycov.stan", data=post_data_Bdrycov_American)

## DIAGNOSTIC(S) FROM PARSER:
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
## Trying to compile a simple C file
## Running /usr/lib64/R/bin/R CMD SHLIB foo.c
## gcc -m64 -I"/usr/include/R" -DNDEBUG -I"/usr/lib64/R/library/Rcpp/include/" -I"/usr/lib64/R/library/RcppEigen/include/Eigen/Dense:1,
## from /usr/lib64/R/library/StanHeaders/include/Stan/math/prim/mat/fun/Eigen.hpp:13,
## from <command-line>:
## /usr/lib64/R/library/RcppEigen/include/Eigen/Core:82:12: fatal error: new: No such file or directory
## #include <new>
## ~~~~~
## compilation terminated.
## make: *** [/usr/lib64/R/etc/Makeconf:167: foo.o] Error 1
##
## SAMPLING FOR MODEL 'Predictive_GP_6dimension_withBS_Bdrycov' NOW (CHAIN 1).
## Chain 1: Iteration: 1 / 200 [ 0%] (Sampling)
## Chain 1: Iteration: 100 / 200 [ 50%] (Sampling)
## Chain 1: Iteration: 200 / 200 [100%] (Sampling)
## Chain 1:
## Chain 1: Elapsed Time: 0 seconds (Warm-up)
## Chain 1: 25.2826 seconds (Sampling)
## Chain 1: 25.2826 seconds (Total)
## Chain 1:

```

##Part 3 Predictions Versus Truth

3-1: Computing Means Standard GP

```
#Computing Mean
y_predict_values_SGP <- extract(pred_gp_SGP,permuted=FALSE)
y_mean_values_SGP <- c(colMeans(y_predict_values_SGP))
y_mean_values_SGP <- y_mean_values_SGP[1:(length(y_mean_values_SGP)-1)]

#Computing Standard Deviation
pred_gp_summary_SGP <- summary(pred_gp_SGP, sd=c("sd"))$summary
pred_gp_sd_SGP <- pred_gp_summary_SGP[, c("sd")]
y_sd_values_SGP <- pred_gp_sd_SGP[1:(length(pred_gp_sd_SGP)-1)]
```

3-2: Computing Means Bdrycov

```
#Computing Mean
y_predict_values_Bdrycov <- extract(pred_gp_Bdrycov,permuted=FALSE)
y_mean_values_Bdrycov <- c(colMeans(y_predict_values_Bdrycov))
y_mean_values_Bdrycov <- y_mean_values_Bdrycov[1:(length(y_mean_values_Bdrycov)-1)]

#Computing Standard Deviation
pred_gp_summary_Bdrycov <- summary(pred_gp_Bdrycov, sd=c("sd"))$summary
pred_gp_sd_Bdrycov <- pred_gp_summary_Bdrycov[, c("sd")]
y_sd_values_Bdrycov <- pred_gp_sd_Bdrycov[1:(length(pred_gp_sd_Bdrycov)-1)]
```

3-3: Plotting Predicted Values against Truth

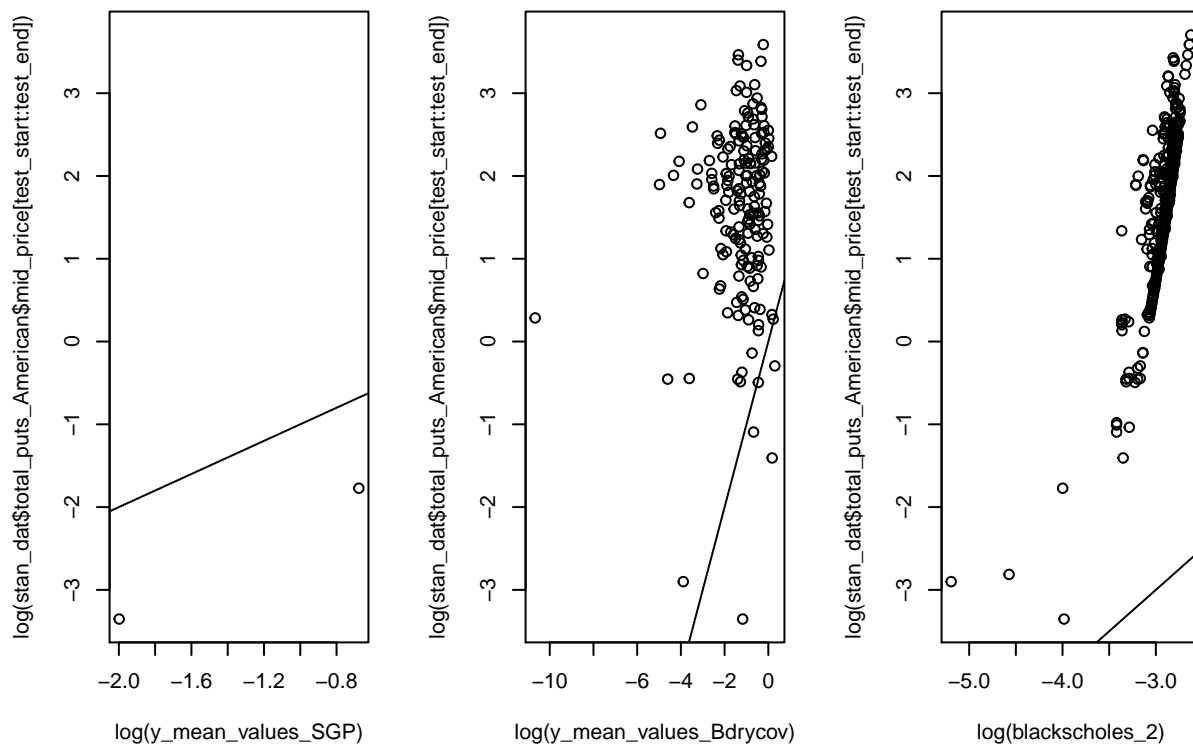
```
par(mfrow=c(1,3))
#Plotting Standard GP
plot(log(y_mean_values_SGP),log(stan_dat$total_puts_American$mid_price[test_start:test_end]))

## Warning in log(y_mean_values_SGP): NaNs produced
abline(0,1)

#Plotting BDrycov
plot(log(y_mean_values_Bdrycov),log(stan_dat$total_puts_American$mid_price[test_start:test_end]))

## Warning in log(y_mean_values_Bdrycov): NaNs produced
abline(0,1)

#Plotting Blackscholes
plot(log(blackscholes_2),log(stan_dat$total_puts_American$mid_price[test_start:test_end]))
abline(0,1)
```



```
#MSE
library('MLmetrics')
```

```
##
## Attaching package: 'MLmetrics'
## The following object is masked from 'package:base':
##
## Recall
```

```
MSE(y_mean_values_SGP,stan_dat$total_puts_American$mid_price[test_start:test_end])
```

```
## [1] 205.4539
```

```
MSE(y_mean_values_Bdrycov,stan_dat$total_puts_American$mid_price[test_start:test_end])
```

```
## [1] 100.1495
```

```
MSE(blackscholes_2,stan_dat$total_puts_American$mid_price[test_start:test_end])
```

```
## [1] 100.3091
```

```
##Part 4 Visualizations
```

4-1: Contour Plots of Forward Price & Strike Price

```
x.grid_1_cont <- as.numeric(stan_dat$total_puts_American$forward_price[test_start:test_end])
x.grid_2_cont <- as.numeric(stan_dat$total_puts_American$strike_price[test_start:test_end])
```

```
dim1 <- seq(min(x.grid_1_cont),max(x.grid_1_cont),length.out = 25)
dim2 <- seq(min(x.grid_2_cont),max(x.grid_2_cont),length.out = 25)
X.grid <- expand.grid(x1 = dim1, x2 = dim2)
```

```
x.grid_3_cont <- as.numeric(rep(mean(stan_dat$total_puts_American$impl_volatility[test_start:test_end]),nrow(X.grid)))
x.grid_4_cont <- as.numeric(rep(mean(stan_dat$total_puts_American$time_to_exp[test_start:test_end]),nrow(X.grid)))
```

```

x.grid_5_cont <- as.numeric(rep(mean(stan_dat$total_puts_American$dividend_yield[test_start:test_end]))
x.grid_6_cont <- as.numeric(rep(mean(stan_dat$total_puts_American$interest_rate[test_start:test_end])),1)

x2_cont <- cbind(X.grid,x.grid_3_cont,x.grid_4_cont,x.grid_5_cont,x.grid_6_cont)

blackscholes_2_cont <- rep(NA,length(x2_cont[,1]))
for (row in 1:nrow(data.frame(x2_cont))){
  blackscholes_2_cont[row] <- as.numeric(blackscholes(-1,S0=x2_cont[row,1],K=x2_cont[row,2],r=x2_cont[r
}]

post_data_cont <- list(theta=c(post_mean_theta_1_Bdrycov,post_mean_theta_2_Bdrycov,post_mean_theta_3_Bd
# post_data

pred_gp_cont <- stan(file="Predictive GP_6dimension_withBS_Bdrycov.stan", data=post_data_cont,iter=200,

## DIAGNOSTIC(S) FROM PARSER:
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
##
##
## SAMPLING FOR MODEL 'Predictive GP_6dimension_withBS_Bdrycov' NOW (CHAIN 1).
## Chain 1: Iteration: 1 / 200 [ 0%] (Sampling)
## Chain 1: Iteration: 100 / 200 [ 50%] (Sampling)
## Chain 1: Iteration: 200 / 200 [100%] (Sampling)
## Chain 1:
## Chain 1: Elapsed Time: 0 seconds (Warm-up)
## Chain 1: 69.8157 seconds (Sampling)
## Chain 1: 69.8157 seconds (Total)
## Chain 1:

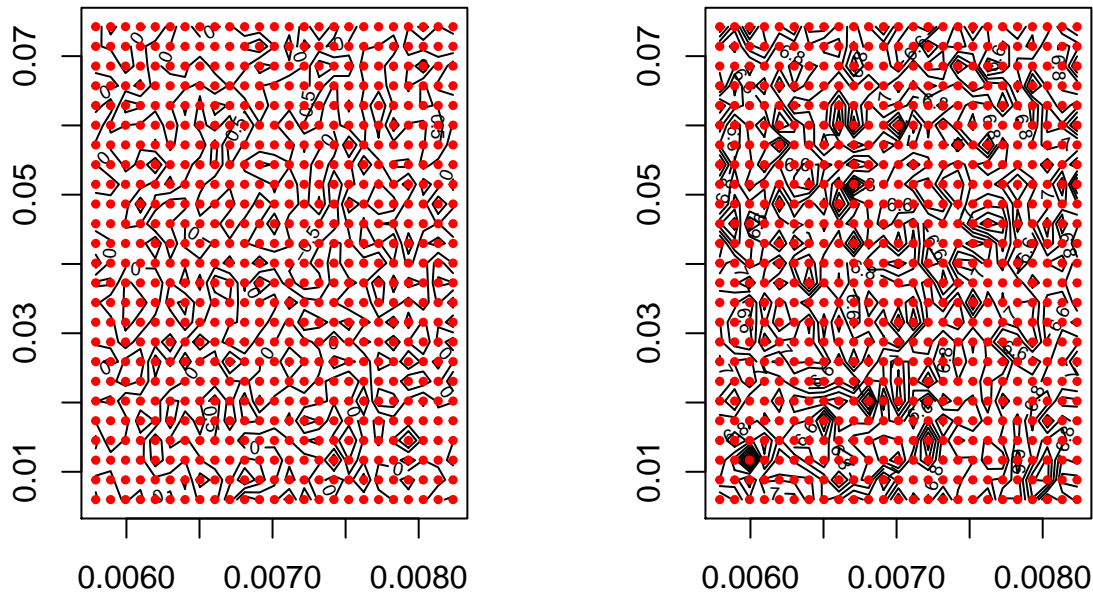
#Computing Mean
y_predict_values_cont <- extract(pred_gp_cont,permuted=FALSE)
y_mean_values_cont <- c(colMeans(y_predict_values_cont))
y_mean_values_cont <- y_mean_values_cont[1:(length(y_mean_values_cont)-1)]

#Computing Standard Deviation
pred_gp_summary_cont <- summary(pred_gp_cont, sd=c("sd"))$summary
pred_gp_sd_cont <- pred_gp_summary_cont[, c("sd")]
y_sd_values_cont <- pred_gp_sd_cont[1:(length(pred_gp_sd_cont)-1)]

par(mfrow = c(1, 2))
#Contour for Predictions aka mean values of predicitions
# x1_grid_cont <- seq(from=min(x.grid_1_cont), to=max(x.grid_1_cont), length.out=length(x.grid_1_cont))
# x2_grid_cont <- seq(from=min(x.grid_2_cont), to=max(x.grid_2_cont), length.out=length(x.grid_2_cont))

contour(dim1, dim2, matrix(y_mean_values_cont, length(dim1), length(dim2)))
points(x2_cont[,1], x2_cont[,2], pch = 19, cex = 0.5, col = "red")
#Contour of Variance
contour(dim1, dim2, matrix(y_sd_values_cont, length(dim1), length(dim2)))
points(x2_cont[,1], x2_cont[,2], pch = 19, cex = 0.5, col = "red")

```



4-2: Contour Plots of Implied Volatility & Time to Expiration

```
x.grid_1_cont <- as.numeric(stan_dat$total_puts_American$impl_volatility[test_start:test_end])
x.grid_2_cont <- as.numeric(stan_dat$total_puts_American$time_to_exp[test_start:test_end])

dim1 <- seq(min(x.grid_1_cont),max(x.grid_1_cont),length.out = 25)
dim2 <- seq(min(x.grid_2_cont),max(x.grid_2_cont),length.out = 25)
X.grid <- expand.grid(x1 = dim1, x2 = dim2)

x.grid_3_cont <- as.numeric(rep(mean(stan_dat$total_puts_American$forward_price[test_start:test_end])),
x.grid_4_cont <- as.numeric(rep(mean(stan_dat$total_puts_American$strike_price[test_start:test_end])),n
x.grid_5_cont <- as.numeric(rep(mean(stan_dat$total_puts_American$dividend_yield[test_start:test_end])),
x.grid_6_cont <- as.numeric(rep(mean(stan_dat$total_puts_American$interest_rate[test_start:test_end])),

x2_cont <- cbind(X.grid,x.grid_3_cont,x.grid_4_cont,x.grid_5_cont,x.grid_6_cont)

blackscholes_2_cont <- rep(NA,length(x2_cont[,1]))
for (row in 1:nrow(data.frame(x2_cont))){
  blackscholes_2_cont[row] <- as.numeric(blackscholes(-1,S0=x2_cont[row,3],K=x2_cont[row,4],r=x2_cont[r
})

post_data_cont <- list(theta=c(post_mean_theta_1_Bdrycov,post_mean_theta_2_Bdrycov,post_mean_theta_3_Bd
# post_data

pred_gp_cont <- stan(file="Predictive GP_6dimension_withBS_Bdrycov.stan", data=post_data_cont,iter=200,

## DIAGNOSTIC(S) FROM PARSER:
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
##
##
## SAMPLING FOR MODEL 'Predictive GP_6dimension_withBS_Bdrycov' NOW (CHAIN 1).
## Chain 1: Iteration: 1 / 200 [ 0%] (Sampling)
## Chain 1: Iteration: 100 / 200 [ 50%] (Sampling)
## Chain 1: Iteration: 200 / 200 [100%] (Sampling)
## Chain 1:
```

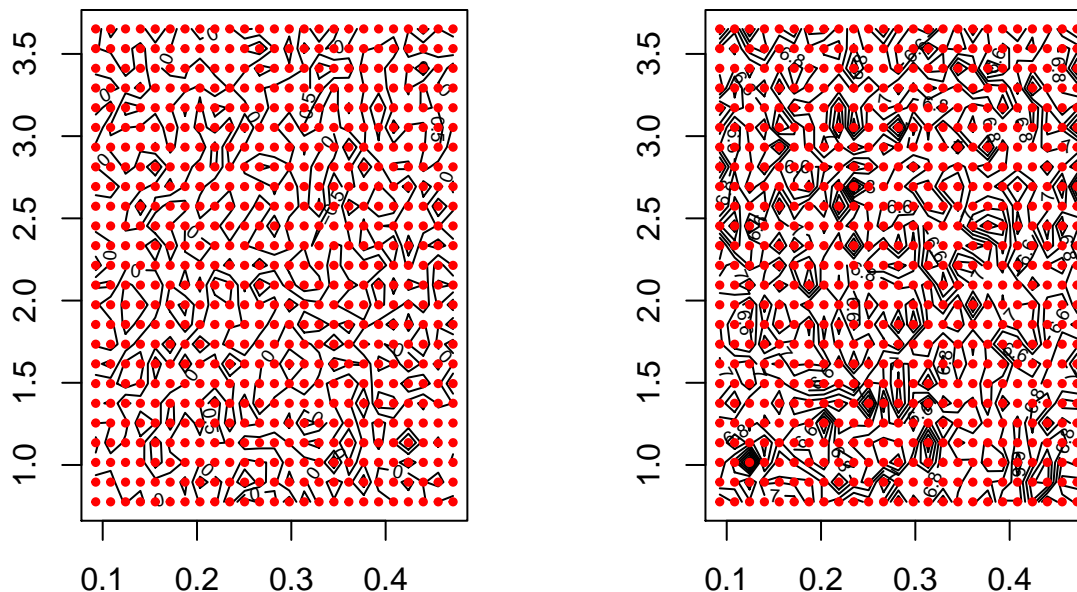
```
## Chain 1: Elapsed Time: 0 seconds (Warm-up)
## Chain 1: 69.5378 seconds (Sampling)
## Chain 1: 69.5378 seconds (Total)
## Chain 1:

#Computing Mean
y_predict_values_cont <- extract(pred_gp_cont, permuted=FALSE)
y_mean_values_cont <- c(colMeans(y_predict_values_cont))
y_mean_values_cont <- y_mean_values_cont[1:(length(y_mean_values_cont)-1)]

#Computing Standard Deviation
pred_gp_summary_cont <- summary(pred_gp_cont, sd=c("sd"))$summary
pred_gp_sd_cont <- pred_gp_summary_cont[, c("sd")]
y_sd_values_cont <- pred_gp_sd_cont[1:(length(pred_gp_sd_cont)-1)]

par(mfrow = c(1, 2))
#Contour for Predictions aka mean values of predicitions
# x1_grid_cont <- seq(from=min(x.grid_1_cont), to=max(x.grid_1_cont), length.out=length(x.grid_1_cont))
# x2_grid_cont <- seq(from=min(x.grid_2_cont), to=max(x.grid_2_cont), length.out=length(x.grid_2_cont))

contour(dim1, dim2, matrix(y_mean_values_cont, length(dim1), length(dim2)))
points(x2_cont[,1], x2_cont[,2], pch = 19, cex = 0.5, col = "red")
#Contour of Variance
contour(dim1, dim2, matrix(y_sd_values_cont, length(dim1), length(dim2)))
points(x2_cont[,1], x2_cont[,2], pch = 19, cex = 0.5, col = "red")
```



4-3: Contour Plots of Interest Rate & Time to Expiration

```
x.grid_1_cont <- as.numeric(stan_dat$total_puts_American$interest_rate[test_start:test_end])
x.grid_2_cont <- as.numeric(stan_dat$total_puts_American$time_to_exp[test_start:test_end])

dim1 <- seq(min(x.grid_1_cont), max(x.grid_1_cont), length.out = 25)
dim2 <- seq(min(x.grid_2_cont), max(x.grid_2_cont), length.out = 25)
X.grid <- expand.grid(x1 = dim1, x2 = dim2)

x.grid_3_cont <- as.numeric(rep(mean(stan_dat$total_puts_American$forward_price[test_start:test_end]),
```



```

x.grid_4_cont <- as.numeric(rep(mean(stan_dat$total_puts_American$strike_price[test_start:test_end])),n)
x.grid_5_cont <- as.numeric(rep(mean(stan_dat$total_puts_American$impl_volatility[test_start:test_end])),n)
x.grid_6_cont <- as.numeric(rep(mean(stan_dat$total_puts_American$dividend_yield[test_start:test_end])),n)

x2_cont <- cbind(X.grid,x.grid_3_cont,x.grid_4_cont,x.grid_5_cont,x.grid_6_cont)

blackscholes_2_cont <- rep(NA,length(x2_cont[,1]))
for (row in 1:nrow(data.frame(x2_cont))){
  blackscholes_2_cont[row] <- as.numeric(blackscholes(-1,S0=x2_cont[row,3],K=x2_cont[row,4],r=x2_cont[row,5],sigma=x2_cont[row,6]))
}

post_data_cont <- list(theta=c(post_mean_theta_1_Bdrycov,post_mean_theta_2_Bdrycov,post_mean_theta_3_Bdrycov),
# post_data

pred_gp_cont <- stan(file="Predictive_GP_6dimension_withBS_Bdrycov.stan", data=post_data_cont,iter=200,

## DIAGNOSTIC(S) FROM PARSER:
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
##
##
## SAMPLING FOR MODEL 'Predictive_GP_6dimension_withBS_Bdrycov' NOW (CHAIN 1).
## Chain 1: Iteration: 1 / 200 [ 0%] (Sampling)
## Chain 1: Iteration: 100 / 200 [ 50%] (Sampling)
## Chain 1: Iteration: 200 / 200 [100%] (Sampling)
## Chain 1:
## Chain 1: Elapsed Time: 0 seconds (Warm-up)
## Chain 1: 70.8091 seconds (Sampling)
## Chain 1: 70.8091 seconds (Total)
## Chain 1:

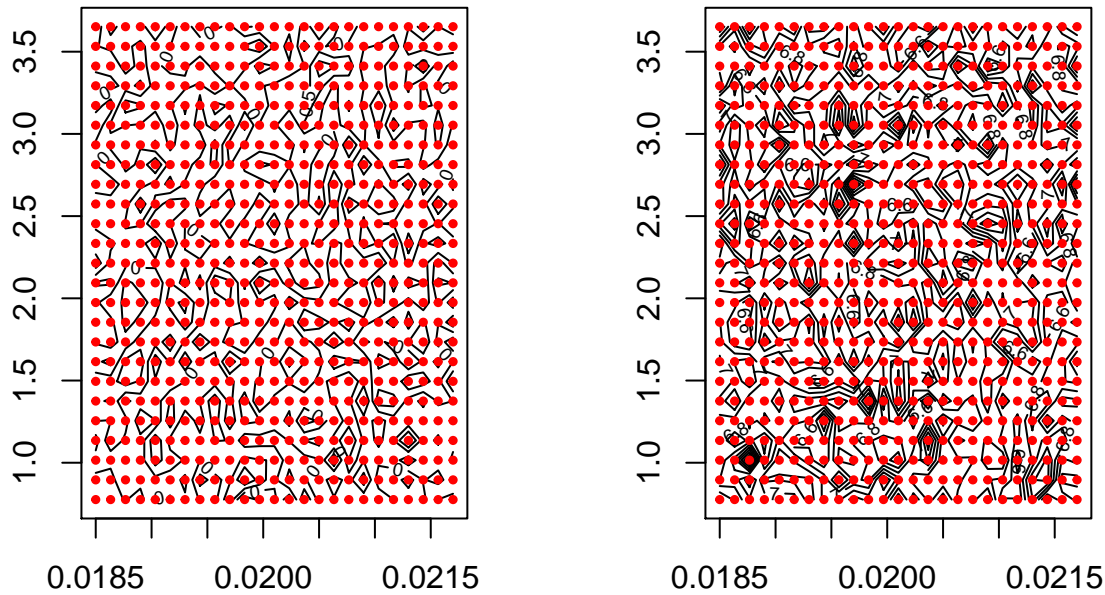
#Computing Mean
y_predict_values_cont <- extract(pred_gp_cont,permuted=FALSE)
y_mean_values_cont <- c(colMeans(y_predict_values_cont))
y_mean_values_cont <- y_mean_values_cont[1:(length(y_mean_values_cont)-1)]

#Computing Standard Deviation
pred_gp_summary_cont <- summary(pred_gp_cont, sd=c("sd"))$summary
pred_gp_sd_cont <- pred_gp_summary_cont[, c("sd")]
y_sd_values_cont <- pred_gp_sd_cont[1:(length(pred_gp_sd_cont)-1)]

par(mfrow = c(1, 2))
#Contour for Predictions aka mean values of predicitons
# x1_grid_cont <- seq(from=min(x.grid_1_cont), to=max(x.grid_1_cont), length.out=length(x.grid_1_cont))
# x2_grid_cont <- seq(from=min(x.grid_2_cont), to=max(x.grid_2_cont), length.out=length(x.grid_2_cont))

contour(dim1, dim2, matrix(y_mean_values_cont, length(dim1), length(dim2)))
points(x2_cont[,1], x2_cont[,2], pch = 19, cex = 0.5, col = "red")
#Contour of Variance
contour(dim1, dim2, matrix(y_sd_values_cont, length(dim1), length(dim2)))
points(x2_cont[,1], x2_cont[,2], pch = 19, cex = 0.5, col = "red")

```

##Part 5: Improving the model by incorporating discrepancy

5-1: Computing Predicted European Option Prices

```
library(rstan)
source("gp.utility.R")

# Fitting GP model
stan_dat_European <- read_rdump('Financial_Data_Put_European.R')

## Parsed with column specification:
## cols(
##   .default = col_double(),
##   date = col_character(),
##   symbol = col_character(),
##   exdate = col_character(),
##   cp_flag = col_character(),
##   ticker = col_character(),
##   exercise_style = col_character()
## )

## See spec(...) for full column specifications.

## Warning: 98350 parsing failures.
##   row      col expected actual
## 142894 6/21/2019 a double FALSE  '~/projects/Independent_Study/spy_spx_(2019.06.01~2019.06.30)_Puts
## 142894 9/20/2019 a double FALSE  '~/projects/Independent_Study/spy_spx_(2019.06.01~2019.06.30)_Puts
## 142894 12/20/2019 a double FALSE  '~/projects/Independent_Study/spy_spx_(2019.06.01~2019.06.30)_Puts
## 142895 6/21/2019 a double FALSE  '~/projects/Independent_Study/spy_spx_(2019.06.01~2019.06.30)_Puts
## 142895 9/20/2019 a double FALSE  '~/projects/Independent_Study/spy_spx_(2019.06.01~2019.06.30)_Puts
## .....
## See problems(...) for more details.

## Warning in blackscholes[row] <- as.numeric(blackscholes[-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes[-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length
```

[illegible]

[illegible]

[illegible]

```

## Warning in blackscholes[row] <- as.numeric(blackscholes[-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

fit_gp_SGP_European <- stan(file="gp-fit-6dimension_withBS.stan", data=stan_dat_European,
                             iter=100, chains=1);

## DIAGNOSTIC(S) FROM PARSER:
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
##
##
## SAMPLING FOR MODEL 'gp-fit-6dimension_withBS' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 0.016684 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 166.84 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:
## Chain 1:
## Chain 1: WARNING: There aren't enough warmup iterations to fit the
## Chain 1:           three stages of adaptation as currently configured.
## Chain 1:           Reducing each adaptation stage to 15%/75%/10% of
## Chain 1:           the given number of warmup iterations:
## Chain 1:           init_buffer = 7
## Chain 1:           adapt_window = 38
## Chain 1:           term_buffer = 5
## Chain 1:
## Chain 1: Iteration:  1 / 100 [  1%] (Warmup)
## Chain 1: Iteration: 10 / 100 [ 10%] (Warmup)
## Chain 1: Iteration: 20 / 100 [ 20%] (Warmup)
## Chain 1: Iteration: 30 / 100 [ 30%] (Warmup)
## Chain 1: Iteration: 40 / 100 [ 40%] (Warmup)
## Chain 1: Iteration: 50 / 100 [ 50%] (Warmup)
## Chain 1: Iteration: 51 / 100 [ 51%] (Sampling)
## Chain 1: Iteration: 60 / 100 [ 60%] (Sampling)
## Chain 1: Iteration: 70 / 100 [ 70%] (Sampling)
## Chain 1: Iteration: 80 / 100 [ 80%] (Sampling)
## Chain 1: Iteration: 90 / 100 [ 90%] (Sampling)
## Chain 1: Iteration: 100 / 100 [100%] (Sampling)
## Chain 1:
## Chain 1: Elapsed Time: 16.6719 seconds (Warm-up)
## Chain 1:           27.9789 seconds (Sampling)
## Chain 1:           44.6508 seconds (Total)
## Chain 1:

## Warning: The largest R-hat is 1.67, indicating chains have not mixed.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#r-hat

## Warning: Bulk Effective Samples Size (ESS) is too low, indicating posterior means and medians may be
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#bulk-ess

## Warning: Tail Effective Samples Size (ESS) is too low, indicating posterior variances and tail quant
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#tail-ess

```

```
print(fit_gp_SGP_European, pars = c('theta','sigma2','gamma2'))
```

```
## Inference for Stan model: gp-fit-6dimension_withBS.
## 1 chains, each with iter=100; warmup=50; thin=1;
## post-warmup draws per chain=50, total post-warmup draws=50.
##
##               mean   se_mean      sd    2.5%    25%    50%    75%
## theta[1]      3.09     0.29     1.84     1.15     1.79     2.62     3.84
## theta[2]      1.88     0.13     0.64     1.03     1.49     1.65     2.29
## theta[3]     17.67     0.37     2.59    12.94    16.26    17.93    18.88
## theta[4]       0.12     0.00     0.02     0.09     0.11     0.12     0.13
## theta[5]    507.67    373.20   1902.90     0.07     0.23     0.51    16.06
## theta[6]      25.29     2.18    14.49     6.06    13.63    22.15    33.64
## sigma2        0.00     0.00     0.00     0.00     0.00     0.00     0.00
## gamma2   831203.51 73980.94 373231.18 321447.70 602426.75 792242.88 971858.93
##               97.5% n_eff Rhat
## theta[1]       7.30     40 0.98
## theta[2]       3.42     23 1.01
## theta[3]      22.99     49 1.02
## theta[4]       0.16     43 0.98
## theta[5]    7687.69     26 1.05
## theta[6]      56.66     44 0.99
## sigma2         0.00     85 1.00
## gamma2   1721184.73     25 0.98
##
## Samples were drawn using NUTS(diag_e) at Thu Mar 26 16:45:54 2020.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).
```

```
sum_gp_SGP_European <- extract(fit_gp_SGP_European,permuted=FALSE)
```

```
# Predicting from GP model
```

```
post_mean_theta_1_SGP <- mean(sum_gp_SGP_European[,1,1]) #theta
post_mean_theta_2_SGP <- mean(sum_gp_SGP_European[,1,2]) #theta
post_mean_theta_3_SGP <- mean(sum_gp_SGP_European[,1,3]) #theta
post_mean_theta_4_SGP <- mean(sum_gp_SGP_European[,1,4]) #theta
post_mean_theta_5_SGP <- mean(sum_gp_SGP_European[,1,5]) #theta
post_mean_theta_6_SGP <- mean(sum_gp_SGP_European[,1,6]) #theta
post_mean_sigma2_SGP <- mean(sum_gp_SGP_European[,1,7]) #sigma2
post_mean_gamma2_SGP <- mean(sum_gp_SGP_European[,1,8]) #gamma2
post_mean_mu_SGP <- stan_dat_European$blackscholes
```

```
# x2 <- as.numeric(unlist(spx_spy_2019_06_30_put_2017_06_500rows_test['strike_price']))
```

```
# x2<- cbind(spy_2013_01_01_2013_01_31_put$strike_price[201:300],spy_2013_01_01_2013_01_31_put$impl_vol
```

```
# x2 <- seq(from=-2,to=2,by=0.01)
```

```
# x2 <- cbind(seq(from=0,to=1,by=0.01),seq(from=0,to=1,by=0.01))
```

```
# test_start <- 323 #06/10
```

```
# test_end <- 559 #06/14
```

```
test_start <- 560 #06/17
```

```

test_end <- 852 #06/20

x.grid_1 <- as.numeric(stan_dat$total_puts_American$forward_price[test_start:test_end])
x.grid_2 <- as.numeric(stan_dat$total_puts_American$strike_price[test_start:test_end])
x.grid_3 <- as.numeric(stan_dat$total_puts_American$simpl_volatility[test_start:test_end])
x.grid_4 <- as.numeric(stan_dat$total_puts_American$time_to_exp[test_start:test_end])
x.grid_5 <- as.numeric(stan_dat$total_puts_American$dividend[test_start:test_end])
x.grid_6 <- as.numeric(stan_dat$total_puts_American$interest_rate[test_start:test_end])
x2 <- cbind(x.grid_1,x.grid_2,x.grid_3,x.grid_4,x.grid_5,x.grid_6)

library('qrmtools')
library('ragtop')
blackscholes_2 <- rep(NA,length(x2[,1]))
for (row in 1:nrow(data.frame(x2))){
  blackscholes_2[row] <- as.numeric(blackscholes(-1,S0=x.grid_1[row],K=x.grid_2[row],r=x.grid_6[row],t=
# blackscholes_2[row] <- Black_Scholes(0,x.grid_1[row],x.grid_6[row],x.grid_3[row],x.grid_2[row],x.gr
})

# X.grid <- expand.grid(x1 = x.grid_1, x2 = x.grid_2)

post_data_Bdrycov_American_disc <- list(theta=c(post_mean_theta_1_SGP,post_mean_theta_2_SGP,post_mean_theta_3_SGP,post_mean_theta_4_SGP,post_mean_theta_5_SGP,post_mean_theta_6_SGP))
# post_data

pred_gp_Bdrycov_disc <- stan(file="Predictive_GP_6dimension_withBS_Bdrycov.stan", data=post_data_Bdrycov_disc)

## DIAGNOSTIC(S) FROM PARSER:
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
##
##
## SAMPLING FOR MODEL 'Predictive_GP_6dimension_withBS_Bdrycov' NOW (CHAIN 1).
## Chain 1: Iteration: 1 / 200 [ 0%] (Sampling)
## Chain 1: Iteration: 100 / 200 [ 50%] (Sampling)
## Chain 1: Iteration: 200 / 200 [100%] (Sampling)
## Chain 1:
## Chain 1: Elapsed Time: 0 seconds (Warm-up)
## Chain 1: 24.8167 seconds (Sampling)
## Chain 1: 24.8167 seconds (Total)
## Chain 1:

#Computing Mean
y_predict_values_Bdrycov_disc <- extract(pred_gp_Bdrycov_disc,permuted=FALSE)
y_mean_values_Bdrycov_disc <- c(colMeans(y_predict_values_Bdrycov_disc))
y_mean_values_Bdrycov_disc <- y_mean_values_Bdrycov_disc[1:(length(y_mean_values_Bdrycov_disc)-1)]

#Computing Standard Deviation
pred_gp_summary_Bdrycov_disc <- summary(pred_gp_Bdrycov_disc, sd=c("sd"))$summary
pred_gp_sd_Bdrycov_disc <- pred_gp_summary_Bdrycov_disc[, c("sd")]
y_sd_values_Bdrycov_disc <- pred_gp_sd_Bdrycov_disc[1:(length(pred_gp_sd_Bdrycov_disc)-1)]

3-3: Plotting Predicted Values against Truth
par(mfrow=c(1,4))
#Plotting Standard GP
plot(log(y_mean_values_SGP),log(stan_dat$total_puts_American$mid_price[test_start:test_end]))

```



```
## Warning in log(y_mean_values_SGP): NaNs produced
```

```
abline(0,1)
```

```
#Plotting Bdrycov
```

```
plot(log(y_mean_values_Bdrycov),log(stan_dat$total_puts_American$mid_price[test_start:test_end]))
```

```
## Warning in log(y_mean_values_Bdrycov): NaNs produced
```

```
abline(0,1)
```

```
#Plotting Blackscholes
```

```
plot(log(blackscholes_2),log(stan_dat$total_puts_American$mid_price[test_start:test_end]))
```

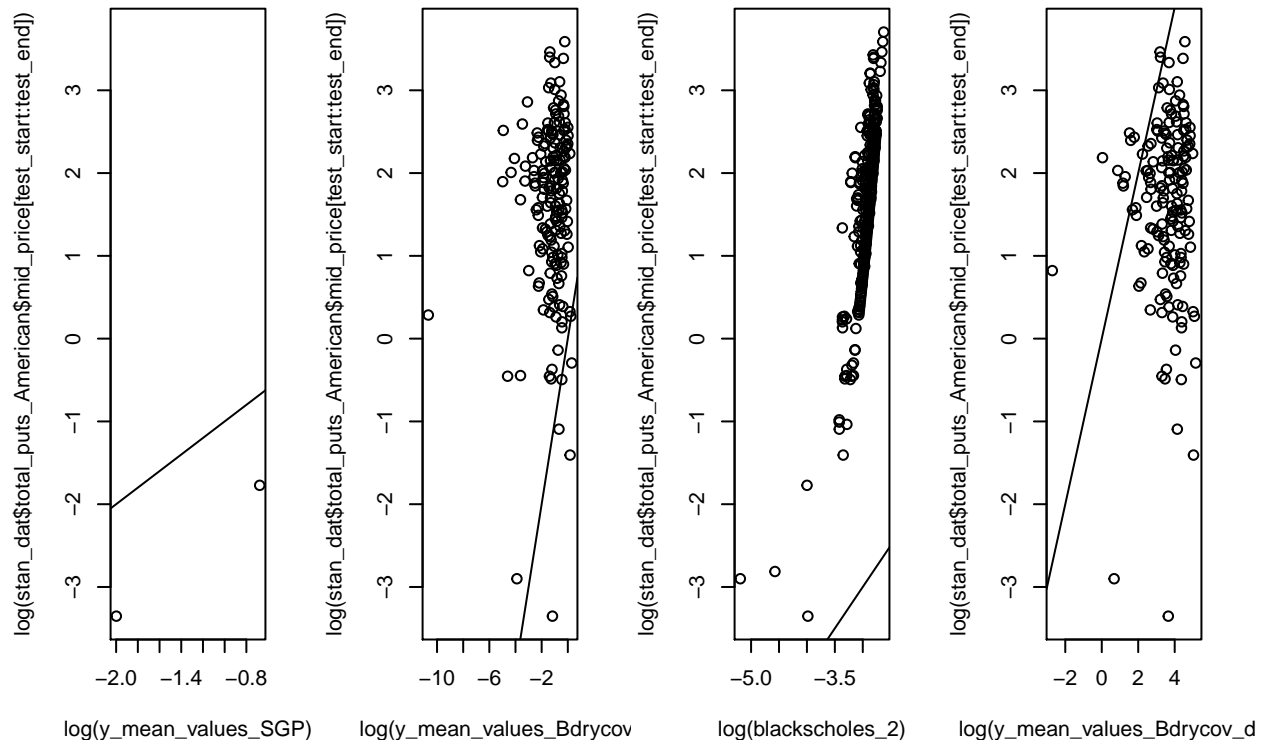
```
abline(0,1)
```

```
#Plotting Blackscholes
```

```
plot(log(y_mean_values_Bdrycov_disc),log(stan_dat$total_puts_American$mid_price[test_start:test_end]))
```

```
## Warning in log(y_mean_values_Bdrycov_disc): NaNs produced
```

```
abline(0,1)
```



```
#MSE
```

```
library('MLmetrics')
```

```
MSE(y_mean_values_SGP,stan_dat$total_puts_American$mid_price[test_start:test_end])
```

```
## [1] 205.4539
```

```
MSE(y_mean_values_Bdrycov,stan_dat$total_puts_American$mid_price[test_start:test_end])
```

```
## [1] 100.1495
```

```
MSE(blackscholes_2,stan_dat$total_puts_American$mid_price[test_start:test_end])
```

```
## [1] 100.3091
MSE(y_mean_values_Bdrycov_disc,stan_dat$total_puts_American$mid_price[test_start:test_end])

## [1] 4205.734
MY EXPERIMENT=====
library(rstan)
source("gp.utility.R")

# Fitting GP model
stan_dat_European_American<- read_rdump('Financial_Data_Put_European_American.R')

## Parsed with column specification:
## cols(
##   .default = col_double(),
##   date = col_character(),
##   symbol = col_character(),
##   exdate = col_character(),
##   cp_flag = col_character(),
##   ticker = col_character(),
##   exercise_style = col_character()
## )
## See spec(...) for full column specifications.
## Warning: 98350 parsing failures.
##   row      col expected actual
## 142894 6/21/2019  a double  FALSE '~/projects/Independent_Study/spy_spx_(2019.06.01~2019.06.30)_Puts
## 142894 9/20/2019  a double  FALSE '~/projects/Independent_Study/spy_spx_(2019.06.01~2019.06.30)_Puts
## 142894 12/20/2019 a double  FALSE '~/projects/Independent_Study/spy_spx_(2019.06.01~2019.06.30)_Puts
## 142895 6/21/2019  a double  FALSE '~/projects/Independent_Study/spy_spx_(2019.06.01~2019.06.30)_Puts
## 142895 9/20/2019  a double  FALSE '~/projects/Independent_Study/spy_spx_(2019.06.01~2019.06.30)_Puts
## .....
## See problems(...) for more details.

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length
```


[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

## Warning in blackscholes[row] <- as.numeric(blackscholes(-1, S0 = x_1[row], :
## number of items to replace is not a multiple of replacement length

fit_gp_SGP_European_American <- stan(file="gp-fit-6dimension_withBS.stan", data=stan_dat_European_Ameri
iter=100, chains=1);

```

```

## DIAGNOSTIC(S) FROM PARSER:
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
##
##
## SAMPLING FOR MODEL 'gp-fit-6dimension_withBS' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 0.090324 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 903.24 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:
## Chain 1:
## Chain 1: WARNING: There aren't enough warmup iterations to fit the
## Chain 1:           three stages of adaptation as currently configured.
## Chain 1:           Reducing each adaptation stage to 15%/75%/10% of
## Chain 1:           the given number of warmup iterations:
## Chain 1:           init_buffer = 7
## Chain 1:           adapt_window = 38
## Chain 1:           term_buffer = 5
## Chain 1:
## Chain 1: Iteration:  1 / 100 [  1%] (Warmup)
## Chain 1: Iteration: 10 / 100 [ 10%] (Warmup)
## Chain 1: Iteration: 20 / 100 [ 20%] (Warmup)
## Chain 1: Iteration: 30 / 100 [ 30%] (Warmup)
## Chain 1: Iteration: 40 / 100 [ 40%] (Warmup)
## Chain 1: Iteration: 50 / 100 [ 50%] (Warmup)
## Chain 1: Iteration: 51 / 100 [ 51%] (Sampling)
## Chain 1: Iteration: 60 / 100 [ 60%] (Sampling)
## Chain 1: Iteration: 70 / 100 [ 70%] (Sampling)
## Chain 1: Iteration: 80 / 100 [ 80%] (Sampling)
## Chain 1: Iteration: 90 / 100 [ 90%] (Sampling)
## Chain 1: Iteration: 100 / 100 [100%] (Sampling)
## Chain 1:
## Chain 1: Elapsed Time: 167.294 seconds (Warm-up)
## Chain 1:           146.327 seconds (Sampling)
## Chain 1:           313.62 seconds (Total)
## Chain 1:

## Warning: The largest R-hat is 1.09, indicating chains have not mixed.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#r-hat

## Warning: Bulk Effective Samples Size (ESS) is too low, indicating posterior means and medians may be
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#bulk-ess

## Warning: Tail Effective Samples Size (ESS) is too low, indicating posterior variances and tail quant.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#tail-ess
print(fit_gp_SGP_European_American, pars = c('theta', 'sigma2', 'gamma2'))

## Inference for Stan model: gp-fit-6dimension_withBS.
## 1 chains, each with iter=100; warmup=50; thin=1;
## post-warmup draws per chain=50, total post-warmup draws=50.
##
##               mean  se_mean      sd    2.5%    25%    50%    75%

```

```
## theta[1]      2.54      0.35      1.76      0.83      1.67      2.00      2.71
## theta[2]      3.63      0.24      1.22      2.17      2.60      3.37      4.44
## theta[3]     13.55      0.16      1.20     11.80     12.56     13.61     14.00
## theta[4]      0.20      0.00      0.02      0.16      0.18      0.19      0.21
## theta[5]    1884.90    355.67    1714.03     23.66     702.13    1432.94    2404.81
## theta[6]      5.64      0.45      2.49      1.34      3.70      5.95      6.79
## sigma2        0.00      0.00      0.00      0.00      0.00      0.00      0.00
## gamma2    563096.03  61134.86  226781.11  263875.71  368746.68  498783.76  746708.26
##
##          97.5% n_eff Rhat
## theta[1]      8.29     25 1.07
## theta[2]      6.15     25 0.99
## theta[3]     16.00     59 0.99
## theta[4]      0.23     30 1.04
## theta[5]    5216.95     23 1.01
## theta[6]     10.54     31 0.98
## sigma2         0.00     52 1.00
## gamma2    969048.96     14 1.06
##
## Samples were drawn using NUTS(diag_e) at Thu Mar 26 16:51:39 2020.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).
```

```
sum_gp_SGP_European_American <- extract(fit_gp_SGP_European_American,permuted=FALSE)
```

```
# Predicting from GP model
```

```
post_mean_theta_1_SGP <- mean(sum_gp_SGP_European_American[,1,1]) #theta
post_mean_theta_2_SGP <- mean(sum_gp_SGP_European_American[,1,2]) #theta
post_mean_theta_3_SGP <- mean(sum_gp_SGP_European_American[,1,3]) #theta
post_mean_theta_4_SGP <- mean(sum_gp_SGP_European_American[,1,4]) #theta
post_mean_theta_5_SGP <- mean(sum_gp_SGP_European_American[,1,5]) #theta
post_mean_theta_6_SGP <- mean(sum_gp_SGP_European_American[,1,6]) #theta
post_mean_sigma2_SGP <- mean(sum_gp_SGP_European_American[,1,7]) #sigma2
post_mean_gamma2_SGP <- mean(sum_gp_SGP_European_American[,1,8]) #gamma2
post_mean_mu_SGP <- stan_dat_European_American$blackscholes
```

```
# x2 <- as.numeric(unlist(spx_spy_2019_06_30_put_2017_06_500rows_test['strike_price']))
```

```
# x2<- cbind(spy_2013_01_01_2013_01_31_put$strike_price[201:300],spy_2013_01_01_2013_01_31_put$impl_vol,
```

```
# x2 <- seq(from=-2,to=2,by=0.01)
```

```
# x2 <- cbind(seq(from=0,to=1,by=0.01),seq(from=0,to=1,by=0.01))
```

```
# test_start <- 323 #06/10
```

```
# test_end <- 559 #06/14
```

```
test_start <- 560 #06/17
```

```
test_end <- 852 #06/20
```

```
x.grid_1 <- as.numeric(stan_dat$total_puts_American$forward_price[test_start:test_end])
x.grid_2 <- as.numeric(stan_dat$total_puts_American$strike_price[test_start:test_end])
x.grid_3 <- as.numeric(stan_dat$total_puts_American$impl_volatility[test_start:test_end])
x.grid_4 <- as.numeric(stan_dat$total_puts_American$time_to_exp[test_start:test_end])
x.grid_5 <- as.numeric(stan_dat$total_puts_American$dividend[test_start:test_end])
x.grid_6 <- as.numeric(stan_dat$total_puts_American$interest_rate[test_start:test_end])
```

```

x2 <- cbind(x.grid_1,x.grid_2,x.grid_3,x.grid_4,x.grid_5,x.grid_6)

library('qrmtools')
library('ragtop')
blackscholes_2 <- rep(NA,length(x2[,1]))
for (row in 1:nrow(data.frame(x2))){
  blackscholes_2[row] <- as.numeric(blackscholes(-1,S0=x.grid_1[row],K=x.grid_2[row],r=x.grid_6[row],t=
  # blackscholes_2[row] <- Black_Scholes(0,x.grid_1[row],x.grid_6[row],x.grid_3[row],x.grid_2[row],x.gr
})

# X.grid <- expand.grid(x1 = x.grid_1, x2 = x.grid_2)

post_data_Bdrycov_European_American_disc <- list(theta=c(post_mean_theta_1_SGP,post_mean_theta_2_SGP,po
# post_data

pred_gp_Bdrycov_disc <- stan(file="Predictive GP_6dimension_withBS_Bdrycov.stan", data=post_data_Bdrycov

## DIAGNOSTIC(S) FROM PARSER:
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
## Info: Comments beginning with # are deprecated. Please use // in place of # for line comments.
##
##
## SAMPLING FOR MODEL 'Predictive GP_6dimension_withBS_Bdrycov' NOW (CHAIN 1).
## Chain 1: Iteration: 1 / 200 [ 0%] (Sampling)
## Chain 1: Iteration: 100 / 200 [ 50%] (Sampling)
## Chain 1: Iteration: 200 / 200 [100%] (Sampling)
## Chain 1:
## Chain 1: Elapsed Time: 0 seconds (Warm-up)
## Chain 1: 52.4274 seconds (Sampling)
## Chain 1: 52.4274 seconds (Total)
## Chain 1:

#Computing Mean
y_predict_values_Bdrycov_disc <- extract(pred_gp_Bdrycov_disc,permuted=FALSE)
y_mean_values_Bdrycov_disc <- c(colMeans(y_predict_values_Bdrycov_disc))
y_mean_values_Bdrycov_disc <- y_mean_values_Bdrycov_disc[1:(length(y_mean_values_Bdrycov_disc)-1)]

#Computing Standard Deviation
pred_gp_summary_Bdrycov_disc <- summary(pred_gp_Bdrycov_disc, sd=c("sd"))$summary
pred_gp_sd_Bdrycov_disc <- pred_gp_summary_Bdrycov_disc[, c("sd")]
y_sd_values_Bdrycov_disc <- pred_gp_sd_Bdrycov_disc[1:(length(pred_gp_sd_Bdrycov_disc)-1)]

3-3: Plotting Predicted Values against Truth

par(mfrow=c(1,4))
#Plotting Standard GP
plot(y_mean_values_SGP,stan_dat$total_puts_American$mid_price[test_start:test_end])
abline(0,1)

#Plotting BDrycov
plot(y_mean_values_Bdrycov,stan_dat$total_puts_American$mid_price[test_start:test_end])
abline(0,1)

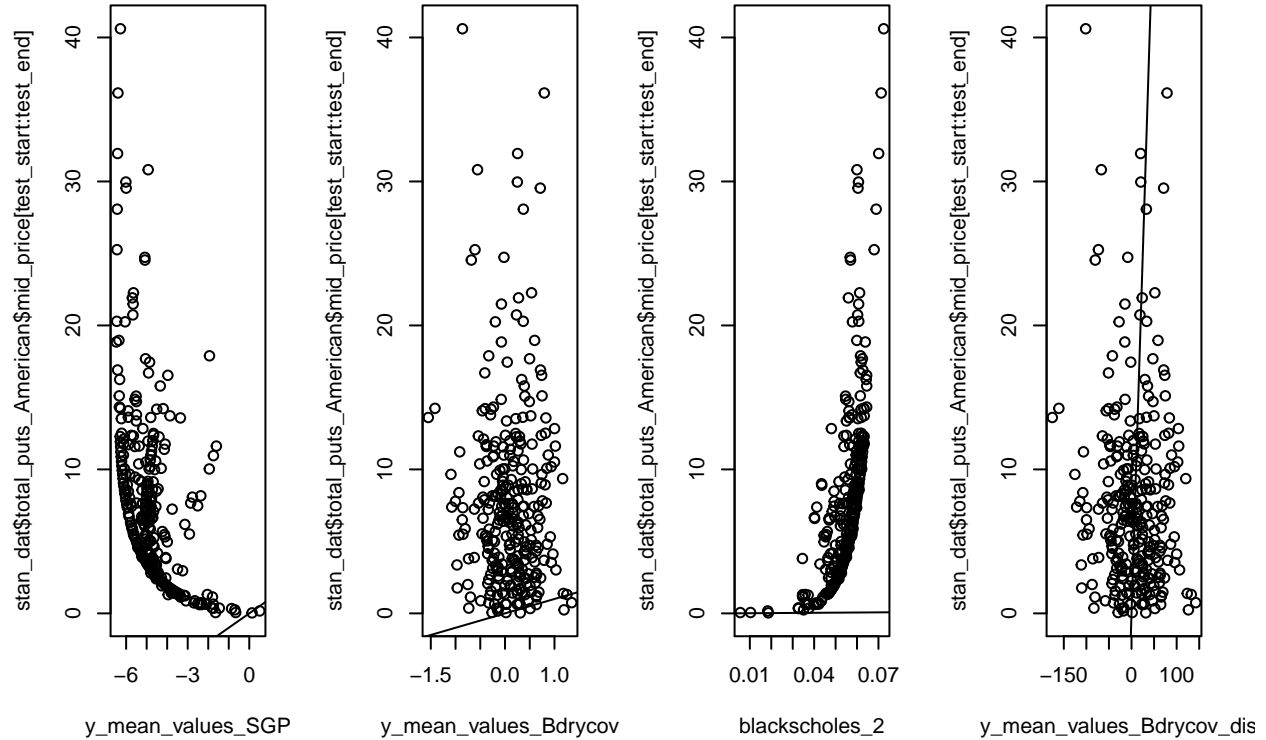
#Plotting Blackscholes
plot(blackscholes_2,stan_dat$total_puts_American$mid_price[test_start:test_end])

```

```
abline(0,1)
```

```
#Plotting Blacksholes
```

```
plot(y_mean_values_Bdrycov_disc,stan_dat$total_puts_American$mid_price[test_start:test_end])
abline(0,1)
```



```
#MSE
```

```
library('MLmetrics')
```

```
MSE(y_mean_values_SGP,stan_dat$total_puts_American$mid_price[test_start:test_end])
```

```
## [1] 205.4539
```

```
MSE(y_mean_values_Bdrycov,stan_dat$total_puts_American$mid_price[test_start:test_end])
```

```
## [1] 100.1495
```

```
MSE(blackscholes_2,stan_dat$total_puts_American$mid_price[test_start:test_end])
```

```
## [1] 100.3091
```

```
MSE(y_mean_values_Bdrycov_disc,stan_dat$total_puts_American$mid_price[test_start:test_end])
```

```
## [1] 2873.863
```