

# 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

```



























[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

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 0.047832 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 478.32 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: 55.1073 seconds (Warm-up)
## Chain 1: 78.6512 seconds (Sampling)
## Chain 1: 133.758 seconds (Total)
## Chain 1:

## Warning: The largest R-hat is 1.3, 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%    97.5%
## theta[1] 200.47   73.70  210.25    0.41   30.20  130.05  317.69  690.28
## theta[2] 292.93   17.45  100.88  118.68  230.23  290.55  343.80  507.77
## theta[3]  18.69    0.52   3.61   12.92   16.28   19.19   20.60   26.19
## theta[4]   0.37    0.01   0.04    0.31    0.34    0.37    0.39    0.44
## theta[5] 6664.58 2482.02 6260.13    4.28  439.46 7255.27 11117.79 18626.43
## theta[6]   0.12    0.02   0.15    0.02    0.06    0.09    0.13    0.36
## sigma2    0.00    0.00   0.00    0.00    0.00    0.00    0.00    0.00
## gamma2  4363.56  508.21 2500.26 1646.72 2503.39 3513.27  5447.10 10670.31
##          n_eff Rhat
## theta[1]     8 1.22
## theta[2]    33 0.99
## theta[3]    48 0.98
## theta[4]    42 1.02
## theta[5]     6 1.28
## theta[6]    61 0.99
## sigma2     76 1.00
## gamma2     24 1.07
##
## Samples were drawn using NUTS(diag_e) at Thu Mar 26 00:44:27 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
test_end <- 559

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, w

## 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/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:           11.7299 seconds (Sampling)
## Chain 1:           11.7299 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
## 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], :

```





[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

## 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_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 0.282034 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 2820.34 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: 1568.11 seconds (Warm-up)
## Chain 1:           1828.83 seconds (Sampling)
## Chain 1:           3396.94 seconds (Total)
## Chain 1:

## Warning: The largest R-hat is 1.56, 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] 3.248720e+08 1.889355e+08 1.098657e+09 0.20    3.10   199.42
## theta[2] 2.368161e+07 1.740719e+07 1.222489e+08 0.11    1.48   341.53
## theta[3] 8.514810e+18 7.611885e+18 4.586875e+19 0.11   21.71  11964.29
## theta[4] 1.741093e+17 1.535311e+17 1.100908e+18 0.17   17.18  29380.15
## theta[5] 6.853307e+22 6.561068e+22 4.845506e+23 0.66 3803.26 5831333.68

```

```
## theta[6] 2.346212e+07 1.848524e+07 8.374132e+07 0.18 29.86 1512.20
## sigma2 7.120000e+01 1.034000e+01 6.672000e+01 0.01 2.13 53.34
## gamma2 7.794000e+01 1.091000e+01 6.852000e+01 0.02 3.92 102.37
##          75%          97.5% n_eff Rhat
## theta[1] 3.297706e+07 2.887776e+09 34 1.07
## theta[2] 3.469258e+04 1.924948e+08 49 0.99
## theta[3] 2.014991e+13 6.131152e+19 36 0.99
## theta[4] 1.243073e+09 7.233774e+17 51 1.01
## theta[5] 8.668018e+14 1.183519e+20 55 1.00
## theta[6] 3.818179e+05 2.738401e+08 21 1.02
## sigma2 1.383700e+02 1.595200e+02 42 0.99
## gamma2 1.392300e+02 1.671100e+02 39 0.99
##
## Samples were drawn using NUTS(diag_e) at Thu Mar 26 01:47:16 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),sigma2=post_mean_sigma2_Bdrycov,gamma2=post_mean_gamma2_Bdrycov,mu=post_mean_mu_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/Eigen/include/Eigen" -fopenmp -std=c++11 -c /usr/lib64/R/library/RcppEigen/src/Eigen/Dense.cpp
```

```
## 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 '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:           27.2542 seconds (Sampling)
## Chain 1:           27.2542 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)

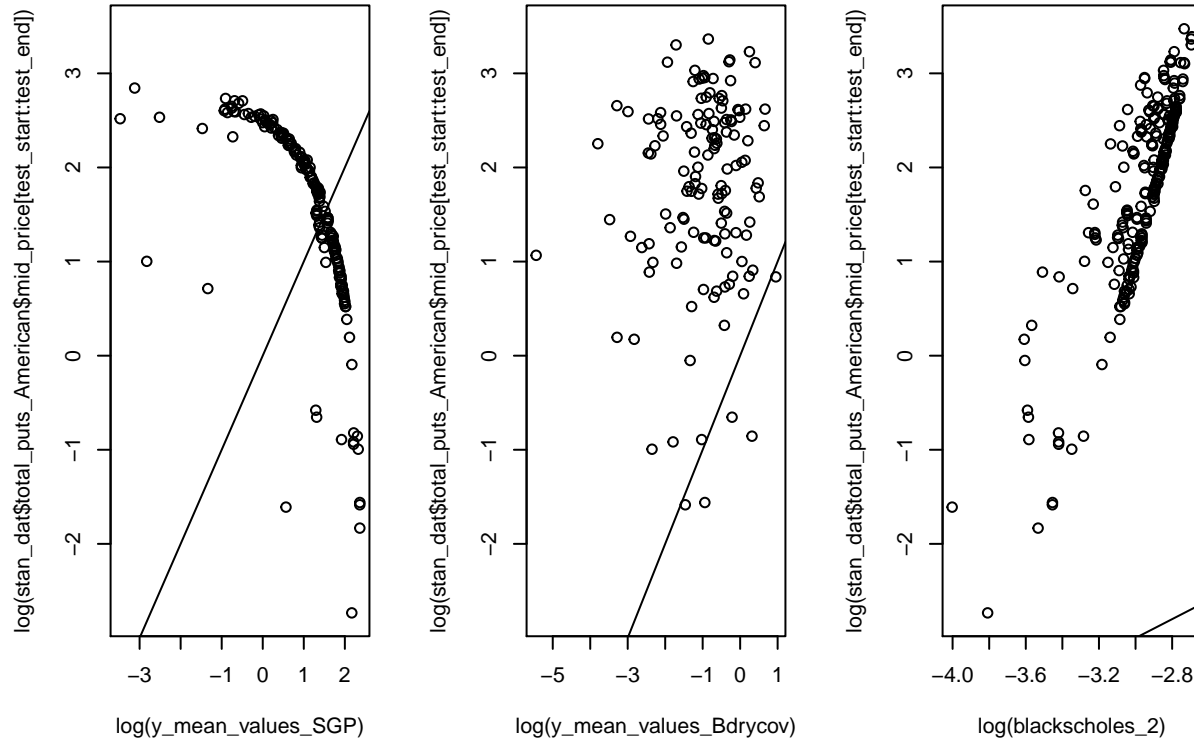
#Plotting Bdrycov
plot(log(y_mean_values_Bdrycov),log(stan_dat$total_puts_American$mid_price[test_start:test_end]))

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

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

```
#Plotting Blacksholes
```

```
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] 116.9114
```

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

```
## [1] 106.3541
```

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

```
## [1] 107.2895
```

```
##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$total_puts_American$simpl_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]),nrow(X.grid)))
x.grid_6_cont <- as.numeric(rep(mean(stan_dat$total_puts_American$interest_rate[test_start:test_end]),nrow(X.grid)))

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[row,3],sigma=x2_cont[row,4],q=x2_cont[row,5]))
}

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: 92.8308 seconds (Sampling)
## Chain 1: 92.8308 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)]

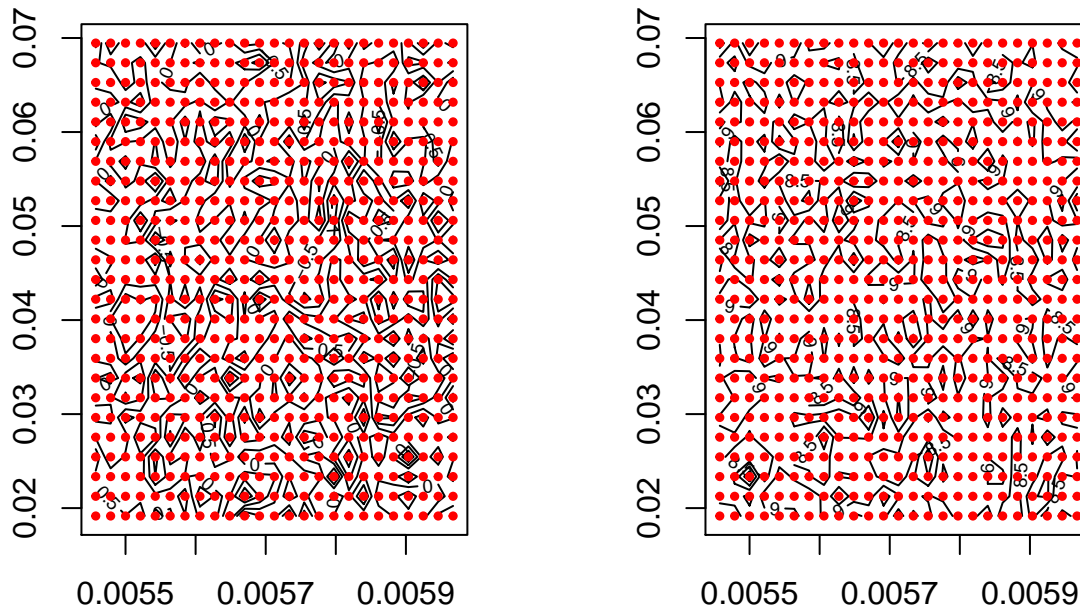
#Computing Standard Deviation
pred_gp_summary_cont <- summary(pred_gp_cont, sd=c("sd"))$summary
pred_gp_sd_cont <- pred_gp_summary_cont[, c("sd")]
y_sd_values_cont <- pred_gp_sd_cont[1:(length(pred_gp_sd_cont)-1)]

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])),n)
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])),n)
x.grid_6_cont <- as.numeric(rep(mean(stan_dat$total_puts_American$interest_rate[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: 91.6661 seconds (Sampling)
## Chain 1: 91.6661 seconds (Total)
## Chain 1:

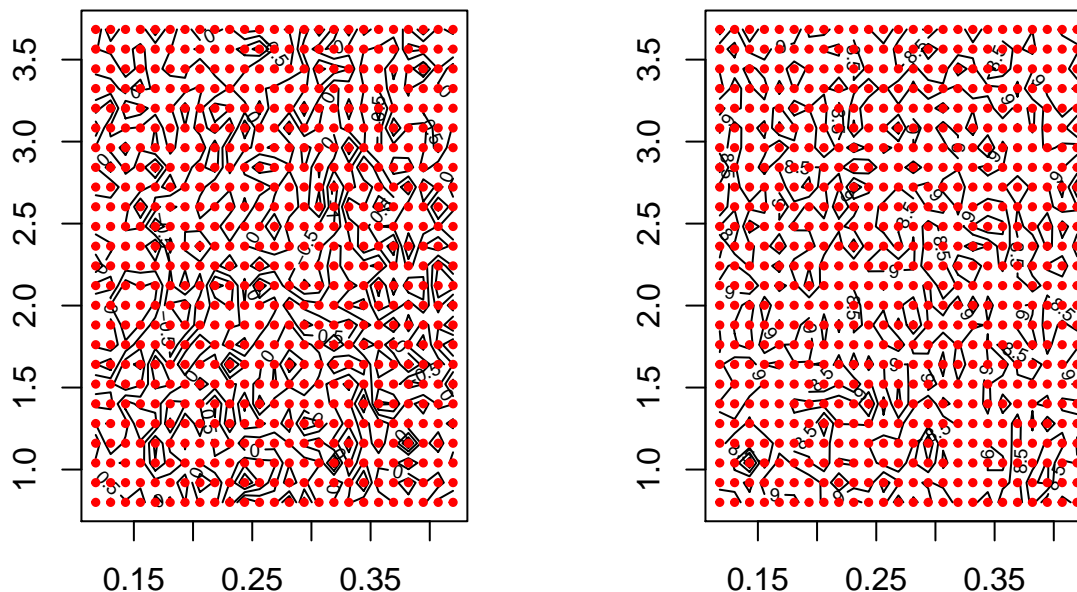
#Computing Mean
y_predict_values_cont <- extract(pred_gp_cont, permuted=FALSE)
y_mean_values_cont <- c(colMeans(y_predict_values_cont))
y_mean_values_cont <- y_mean_values_cont[1:(length(y_mean_values_cont)-1)]

#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])),n)
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: 87.7925 seconds (Sampling)
## Chain 1: 87.7925 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)]

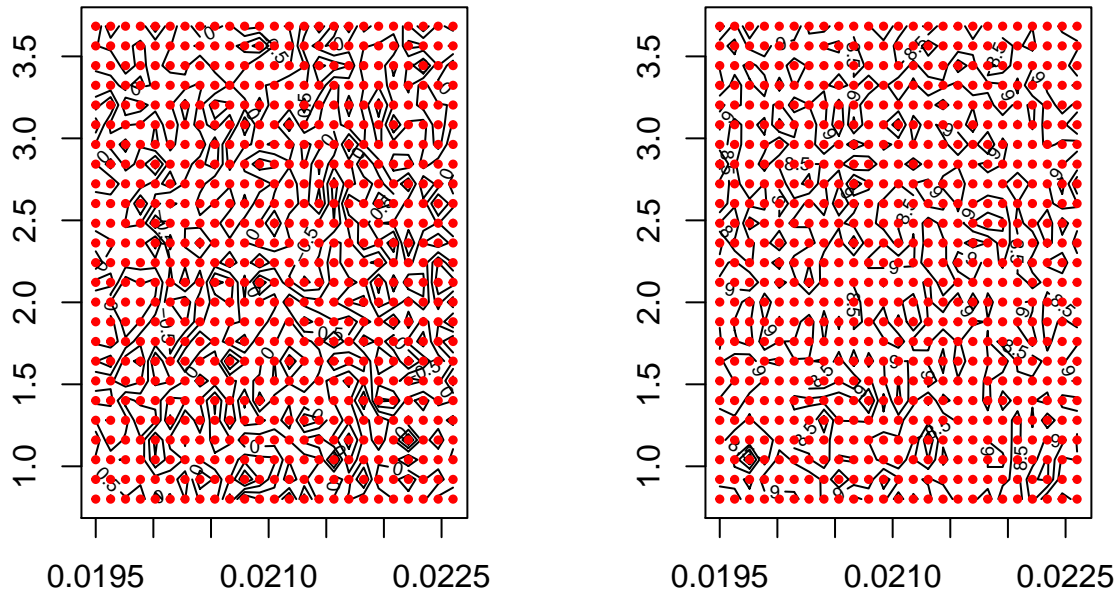
#Computing Standard Deviation
pred_gp_summary_cont <- summary(pred_gp_cont, sd=c("sd"))$summary
pred_gp_sd_cont <- pred_gp_summary_cont[, c("sd")]
y_sd_values_cont <- pred_gp_sd_cont[1:(length(pred_gp_sd_cont)-1)]

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
```





[illegible]



[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
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 0.085132 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 851.32 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: 211.223 seconds (Warm-up)
## Chain 1:           293.674 seconds (Sampling)
## Chain 1:           504.896 seconds (Total)
## Chain 1:

## Warning: The largest R-hat is 1.4, 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%
## theta[1] 6.470000e+00 2.700000e-01 1.680000e+00   3.91    5.48    6.18
## theta[2] 5.890000e+00 1.800000e-01 1.070000e+00   4.02    5.14    5.84
## theta[3] 1.386000e+01 5.600000e-01 3.370000e+00   8.09   11.28   13.90
## theta[4] 7.700000e-01 1.000000e-02 5.000000e-02    0.66    0.74    0.77
## theta[5] 9.383343e+13 8.997897e+13 6.468205e+14    0.60    2.85   41.76
## theta[6] 2.010000e+00 1.400000e-01 8.600000e-01    0.89    1.38    1.87
## sigma2    0.000000e+00 0.000000e+00 0.000000e+00    0.00    0.00    0.00
## gamma2    2.609551e+05 1.067275e+04 6.035227e+04 175404.02 219920.09 255109.11
##
##              75%          97.5% n_eff Rhat
## theta[1]      6.99 1.079000e+01   39 1.04
## theta[2]      6.66 7.700000e+00   34 1.00
## theta[3]     15.41 2.095000e+01   36 1.01
## theta[4]      0.80 8.600000e-01   35 1.08
## theta[5]    1234.43 9.040243e+13   52 1.00
## theta[6]      2.32 3.480000e+00   37 1.03
## sigma2        0.00 0.000000e+00   41 0.98
## gamma2    286916.99 4.037101e+05   32 1.05
##
## Samples were drawn using NUTS(diag_e) at Thu Mar 26 02:03:25 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
```

```
test_end <- 559
```

```
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_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: 40.7752 seconds (Sampling)
## Chain 1: 40.7752 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)

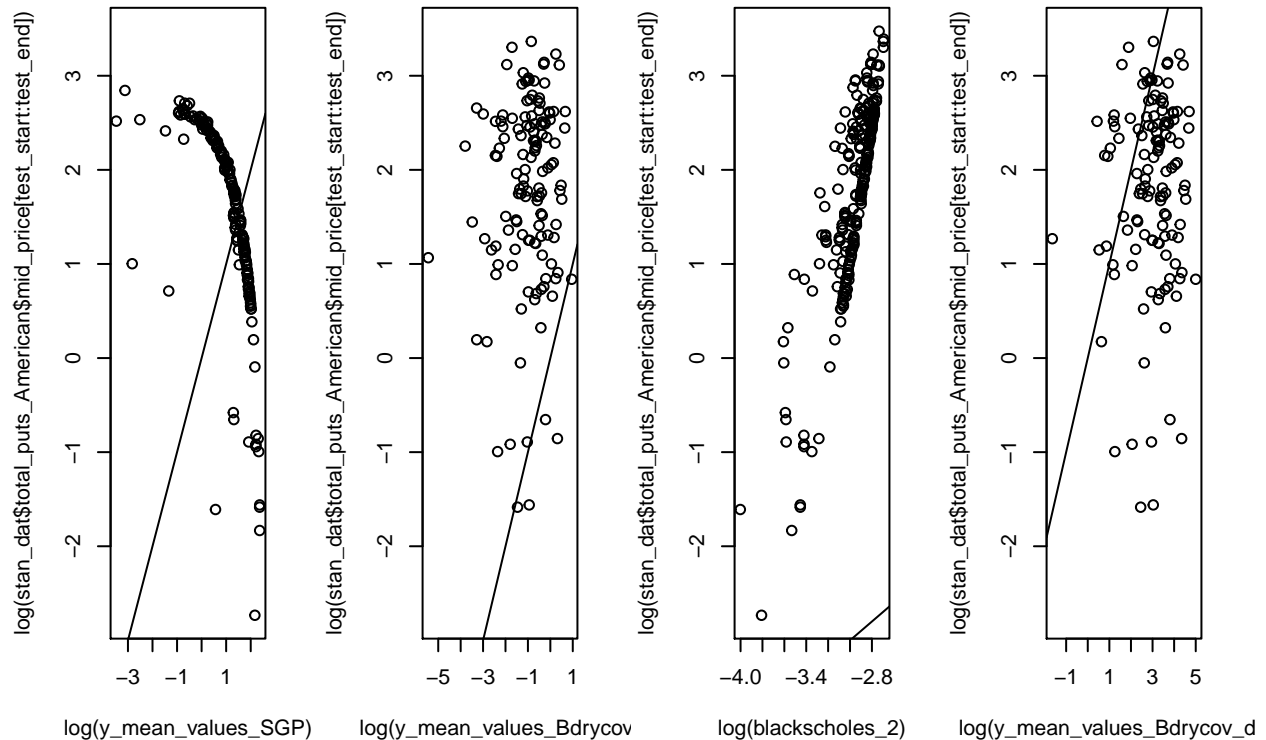
#Plotting Bdrycov
plot(log(y_mean_values_Bdrycov),log(stan_dat$total_puts_American$mid_price[test_start:test_end]))

## 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] 116.9114

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

## [1] 106.3541

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

## [1] 107.2895
```

```

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

## [1] 1530.634

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]

[illegible]

[illegible]

[illegible]

[illegible]



[illegible]

[illegible]



[illegible]

[illegible]

[illegible]



[illegible]

[illegible]

[illegible]





[illegible]

[illegible]

[illegible]



[illegible]

[illegible]

[illegible]

[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

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 0.321333 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 3213.33 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: 236.383 seconds (Warm-up)
## Chain 1: 398.748 seconds (Sampling)
## Chain 1: 635.131 seconds (Total)
## Chain 1:

## Warning: The largest R-hat is 1.21, 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]       7.65   0.26   1.77       4.72       6.80       7.56       8.33
## theta[2]       7.97   0.14   0.87       6.60       7.28       7.86       8.59
## theta[3]      13.05   0.19   1.22      11.17      12.19      12.95      13.85
## theta[4]        0.85   0.01   0.04        0.79        0.83        0.84        0.87
## theta[5]     2212.40 169.44 860.28     771.49    1488.60    2250.15    2730.76
## theta[6]        1.52   0.11   0.69        0.66        1.06        1.32        1.86
## sigma2         0.00   0.00   0.00        0.00        0.00        0.00        0.00
## gamma2    177275.97 8306.99 35505.07 119996.70 150059.73 179267.49 201898.45
##               97.5% n_eff Rhat
## theta[1]      11.55    45 0.98
## theta[2]       9.61    40 1.00
## theta[3]      15.60    41 0.99
## theta[4]        0.92    50 0.98
## theta[5]     3772.61    26 0.99
## theta[6]        3.29    37 0.99
## sigma2         0.00    85 0.98
## gamma2     250740.81    18 0.98
##
## Samples were drawn using NUTS(diag_e) at Thu Mar 26 02:14:56 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
```

```
test_end <- 559
```

```
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, 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: 94.5367 seconds (Sampling)
## Chain 1: 94.5367 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)]
```

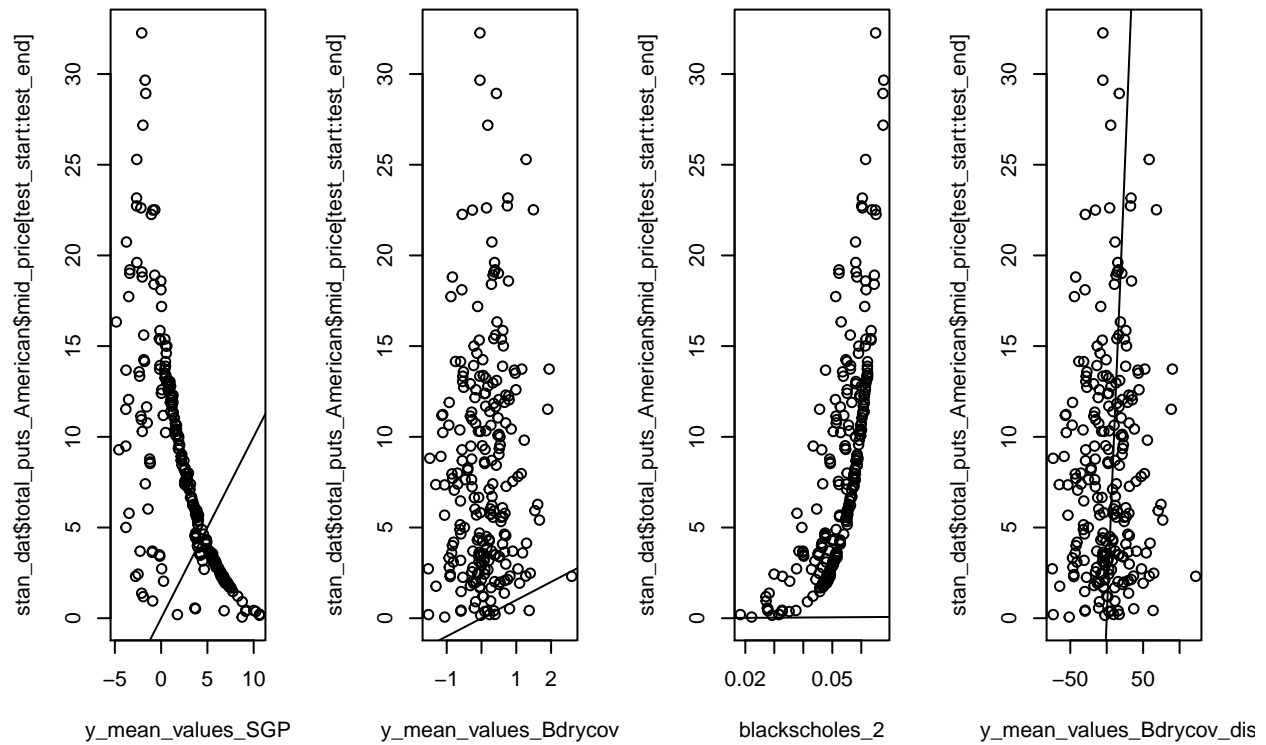
3-3: Plotting Predicted Values against Truth

```
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 Blackscholes
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] 116.9114

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

## [1] 106.3541

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

## [1] 107.2895

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

## [1] 1062.61
```