

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]

[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 0.050865 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 508.65 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: 112.089 seconds (Warm-up)
## Chain 1: 101.339 seconds (Sampling)
## Chain 1: 213.427 seconds (Total)
## Chain 1:

## Warning: The largest R-hat is 1.49, 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]   17.76   10.33   37.09     0.03    0.86    1.81    9.51   132.46
## theta[2]  290.14    9.20   67.81   181.44   239.53   285.80   332.73   414.89
## theta[3]   19.37    0.42    2.99    13.08    17.44    19.47    20.83    24.79
## theta[4]    0.37    0.00    0.04     0.31    0.35    0.38     0.40     0.43
## theta[5] 13664.55  973.32 4840.56 6384.94 10531.29 13729.38 15861.18 25517.83
## theta[6]    0.07    0.03    0.12     0.00    0.01    0.02     0.08     0.31
## sigma2     0.00    0.00    0.00     0.00    0.00    0.00     0.00     0.00
## gamma2   3811.86  198.87 1264.57 2120.68 2843.10 3466.54 4738.46 6287.98
##          n_eff Rhat
## theta[1]    13 1.13
## theta[2]    54 1.05
## theta[3]    51 1.01
## theta[4]    71 0.98
## theta[5]    25 1.11
## theta[6]    19 1.02
## sigma2     85 0.99
## gamma2     40 1.09
##
## Samples were drawn using NUTS(diag_e) at Wed Mar 25 16:58:27 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:           12.3427 seconds (Sampling)
## Chain 1:           12.3427 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]

[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 0.295473 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 2954.73 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: 3737.46 seconds (Warm-up)
## Chain 1:           7289.77 seconds (Sampling)
## Chain 1:           11027.2 seconds (Total)
## Chain 1:

## Warning: There were 34 transitions after warmup that exceeded the maximum treedepth. Increase max_tr
## http://mc-stan.org/misc/warnings.html#maximum-treedepth-exceeded

## Warning: Examine the pairs() plot to diagnose sampling problems

## Warning: The largest R-hat is 1.55, 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%
## theta[1] 4.189108e+151      NaN 2.962138e+152 3.013239e+08 4.156126e+28

```

```
## theta[2] 5.130032e+127 4.911708e+127 3.627468e+128 8.169989e+14 1.562898e+44
## theta[3] 2.928931e+40 2.804291e+40 2.071067e+41 7.000000e-02 2.557000e+02
## theta[4] 1.584579e+77 1.517147e+77 1.120466e+78 1.300000e-01 2.296737e+06
## theta[5] 4.852745e+115 4.646238e+115 3.431409e+116 7.540597e+04 3.846524e+08
## theta[6] 3.443302e+67 3.296773e+67 2.434747e+68 1.731413e+04 8.731309e+18
## sigma2 7.678000e+01 1.006000e+01 6.622000e+01 1.000000e-02 2.970000e+00
## gamma2 6.879000e+01 9.910000e+00 6.505000e+01 2.000000e-02 2.050000e+00
##          50%          75%          97.5% n_eff Rhat
## theta[1] 8.672554e+36 9.503428e+47 4.978483e+147 NaN 1.00
## theta[2] 5.274972e+68 3.061562e+89 4.801216e+123 55 1.00
## theta[3] 2.209660e+07 6.406221e+13 2.276488e+32 55 1.00
## theta[4] 7.939309e+19 1.055313e+32 4.109929e+62 55 1.00
## theta[5] 1.927789e+13 6.621568e+29 4.058353e+102 55 1.00
## theta[6] 5.207723e+26 3.072951e+36 1.816784e+64 55 1.00
## sigma2 8.714000e+01 1.393800e+02 1.581000e+02 43 0.99
## gamma2 5.307000e+01 1.367600e+02 1.538500e+02 43 0.99
##
## Samples were drawn using NUTS(diag_e) at Wed Mar 25 20:08:08 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/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_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: 29.9627 seconds (Sampling)
## Chain 1: 29.9627 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
```

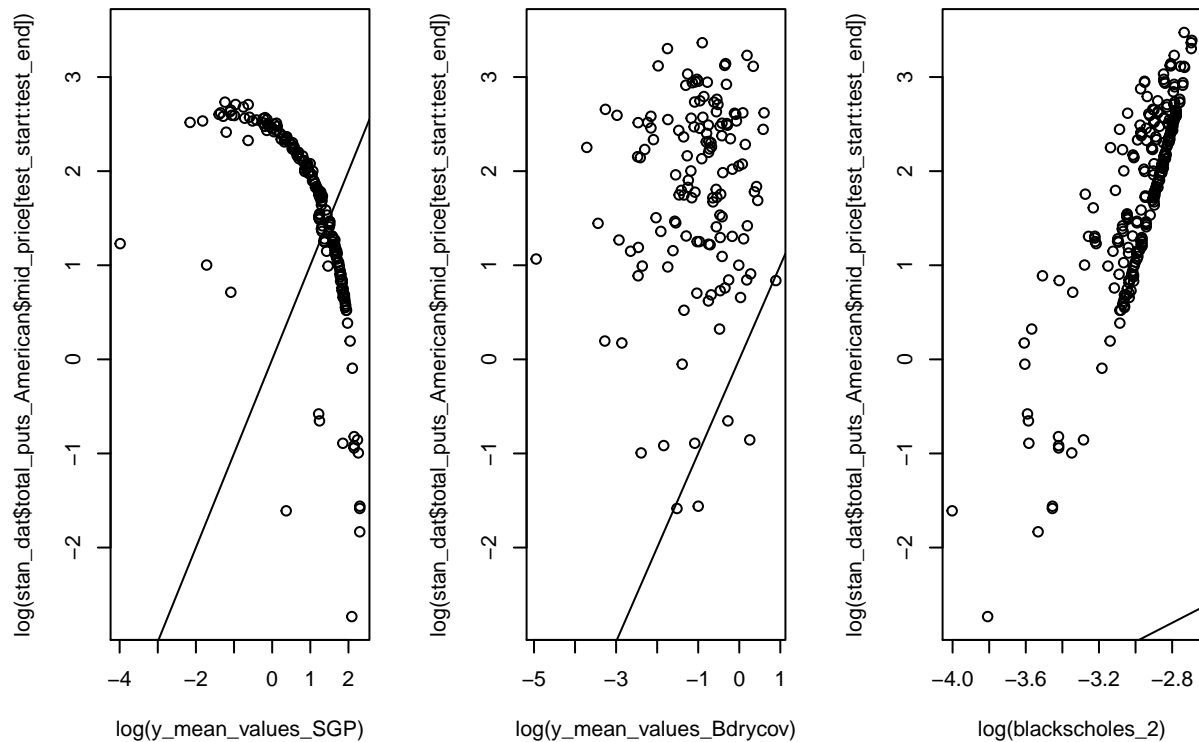
```
abline(0,1)

#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] 115.4352
MSE(y_mean_values_Bdrycov,stan_dat$total_puts_American$mid_price[test_start:test_end])

## [1] 106.3851
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$total_puts_American$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]),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: 93.0917 seconds (Sampling)
## Chain 1: 93.0917 seconds (Total)
## Chain 1:

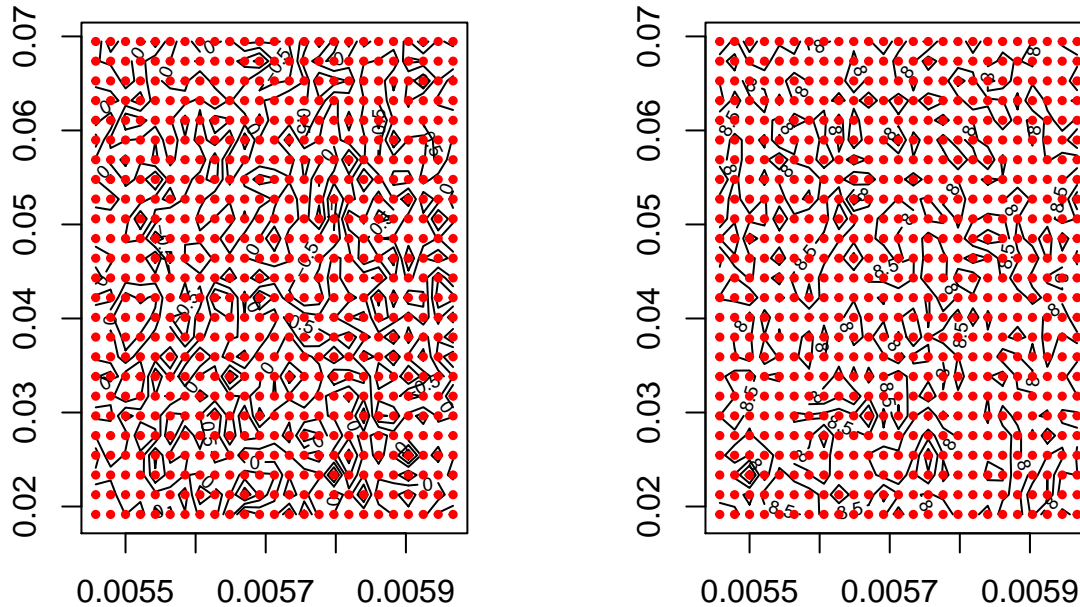
#Computing Mean
y_predict_values_cont <- extract(pred_gp_cont,permuted=FALSE)
y_mean_values_cont <- c(colMeans(y_predict_values_cont))
y_mean_values_cont <- y_mean_values_cont[1:(length(y_mean_values_cont)-1)]

#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],
  })

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: 93.2494 seconds (Sampling)
## Chain 1: 93.2494 seconds (Total)
## Chain 1:

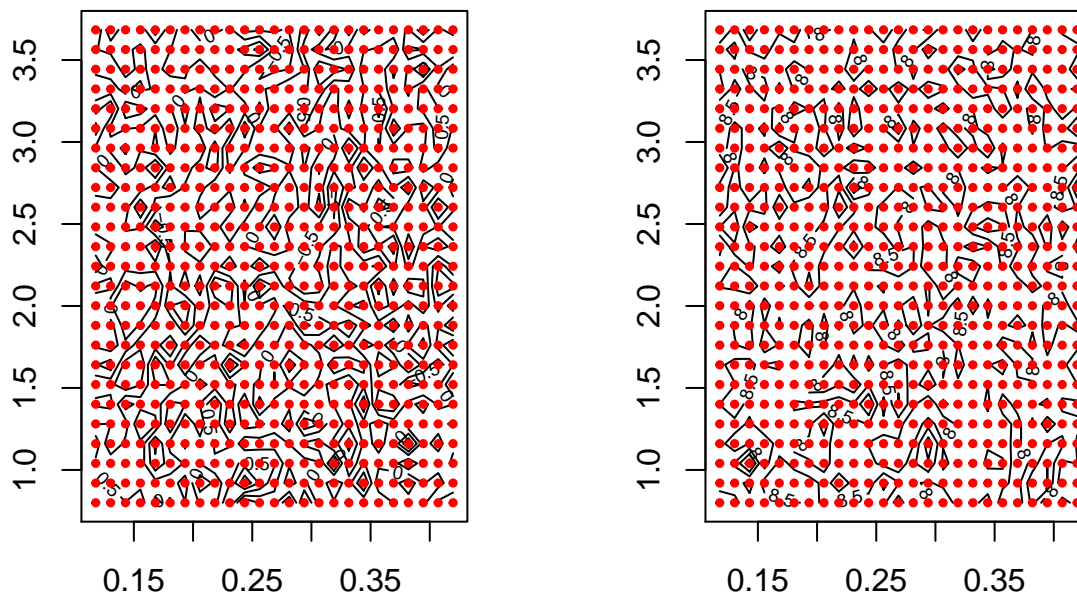
#Computing Mean
y_predict_values_cont <- extract(pred_gp_cont, permuted=FALSE)
y_mean_values_cont <- c(colMeans(y_predict_values_cont))
y_mean_values_cont <- y_mean_values_cont[1:(length(y_mean_values_cont)-1)]

#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 PARSE:
## 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: 92.1401 seconds (Sampling)
## Chain 1: 92.1401 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)]

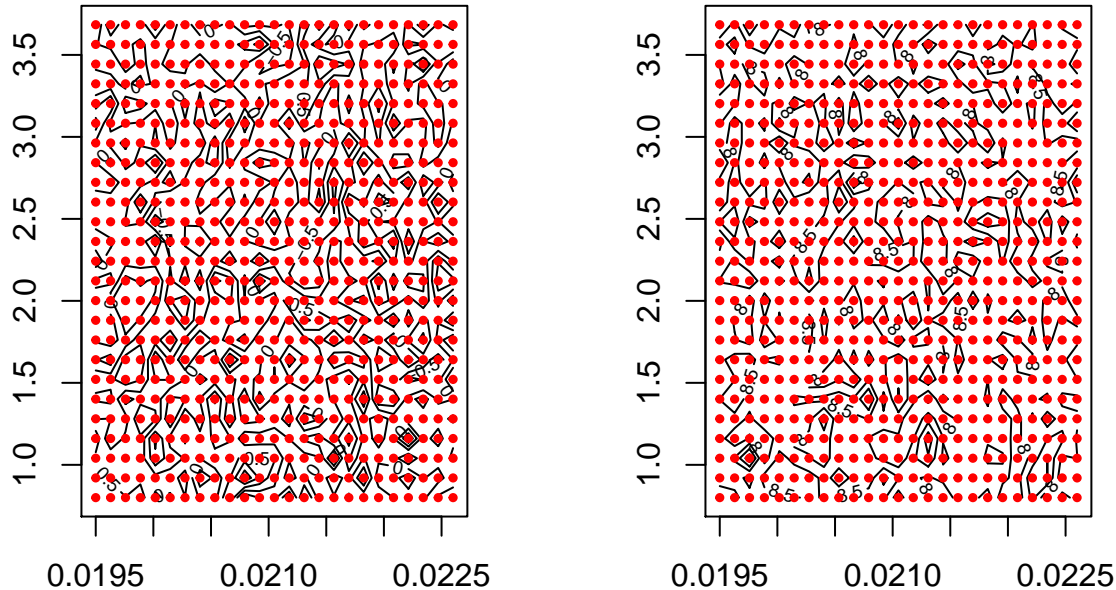
par(mfrow = c(1, 2))
#Contour for Predictions aka mean values of predicitons
# x1_grid_cont <- seq(from=min(x.grid_1_cont), to=max(x.grid_1_cont), length.out=length(x.grid_1_cont))
# x2_grid_cont <- seq(from=min(x.grid_2_cont), to=max(x.grid_2_cont), length.out=length(x.grid_2_cont))

```

```

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
## .....

```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[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
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 0.078273 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 782.73 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: 488.333 seconds (Warm-up)
## Chain 1:           540.176 seconds (Sampling)
## Chain 1:           1028.51 seconds (Total)
## Chain 1:
## Warning: There were 5 transitions after warmup that exceeded the maximum treedepth. Increase max_tre
## http://mc-stan.org/misc/warnings.html#maximum-treedepth-exceeded

## Warning: Examine the pairs() plot to diagnose sampling problems

## Warning: The largest R-hat is 1.2, 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%
## theta[1]	6.570000e+00	2.300000e-01	2.130000e+00	3.500000e+00	5.150000e+00
## theta[2]	5.910000e+00	1.400000e-01	9.600000e-01	4.070000e+00	5.350000e+00
## theta[3]	1.384000e+01	4.100000e-01	2.630000e+00	8.520000e+00	1.214000e+01
## theta[4]	7.700000e-01	1.000000e-02	6.000000e-02	6.700000e-01	7.300000e-01
## theta[5]	4.058892e+109	3.450853e+109	2.554092e+110	4.695805e+54	7.248468e+83
## theta[6]	2.170000e+00	1.800000e-01	1.170000e+00	8.200000e-01	1.210000e+00
## sigma2	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00
## gamma2	2.610477e+05	1.031152e+04	6.661598e+04	1.551891e+05	2.166162e+05

```
##
```

	50%	75%	97.5%	n_eff	Rhat
## theta[1]	6.220000e+00	7.770000e+00	1.085000e+01	85	0.99
## theta[2]	5.920000e+00	6.380000e+00	8.000000e+00	48	0.98
## theta[3]	1.397000e+01	1.576000e+01	1.831000e+01	41	0.98
## theta[4]	7.600000e-01	8.100000e-01	8.900000e-01	53	0.98
## theta[5]	1.349052e+95	6.784663e+95	1.105556e+110	55	1.01
## theta[6]	1.970000e+00	3.030000e+00	4.290000e+00	44	0.99
## sigma2	0.000000e+00	0.000000e+00	0.000000e+00	54	1.00
## gamma2	2.528923e+05	2.998788e+05	3.986715e+05	42	0.98

```
##
```

```
## Samples were drawn using NUTS(diag_e) at Wed Mar 25 20:33:05 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: 44.3622 seconds (Sampling)
## Chain 1: 44.3622 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]))

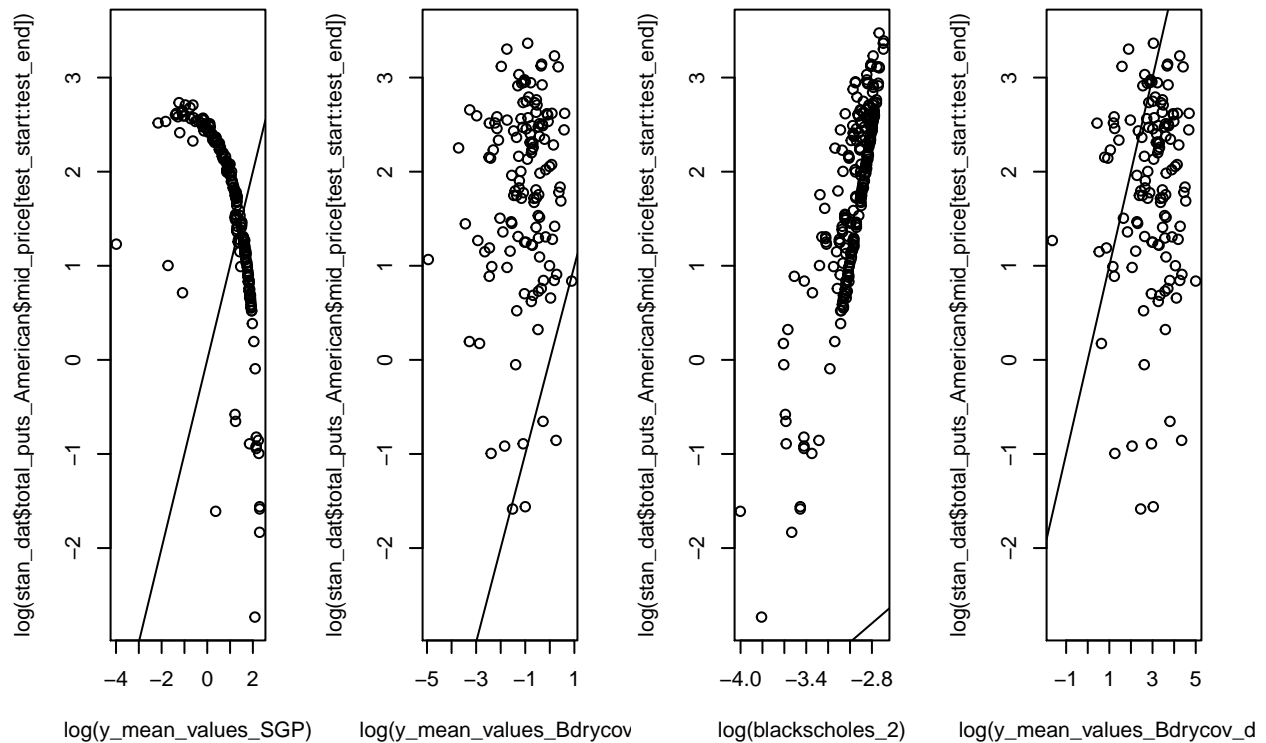
## Warning in log(y_mean_values_Bdrycov): NaNs produced
abline(0,1)

#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] 115.4352
```

```

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

## [1] 106.3851
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] 1531.153
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

```

[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

## 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 0.354563 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 3545.63 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: 359.623 seconds (Warm-up)
## Chain 1:      381.514 seconds (Sampling)
## Chain 1:      741.137 seconds (Total)
## Chain 1:

## Warning: The largest R-hat is 1.29, 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]    8.05    0.29    1.60    5.49    6.79    7.77    9.38
## theta[2]    8.04    0.08    0.65    6.43    7.69    8.09    8.36
## theta[3]   13.12    0.16    1.19   10.99   12.35   13.05   13.89
## theta[4]    0.85    0.00    0.03    0.80    0.83    0.85    0.87
## theta[5]  2383.91  156.17  964.58  964.08  1853.37  2305.09  2792.14
## theta[6]    1.49    0.15    0.63    0.78    1.05    1.31    1.83
## sigma2      0.00    0.00    0.00    0.00    0.00    0.00    0.00
## gamma2  164710.29 3025.20 26511.29 120693.03 151238.61 161983.70 176534.18
##          97.5% n_eff Rhat
## theta[1]   11.26    30 1.03
## theta[2]    9.20    73 0.98
## theta[3]   15.71    53 0.99
## theta[4]    0.91    41 0.99

```



```

## theta[5]    4612.38    38 0.98
## theta[6]      2.94    19 0.98
## sigma2       0.00    85 0.99
## gamma2    220438.31    77 1.00
##
## Samples were drawn using NUTS(diag_e) at Wed Mar 25 20:46:27 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_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: 92.6829 seconds (Sampling)
## Chain 1: 92.6829 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)]
```

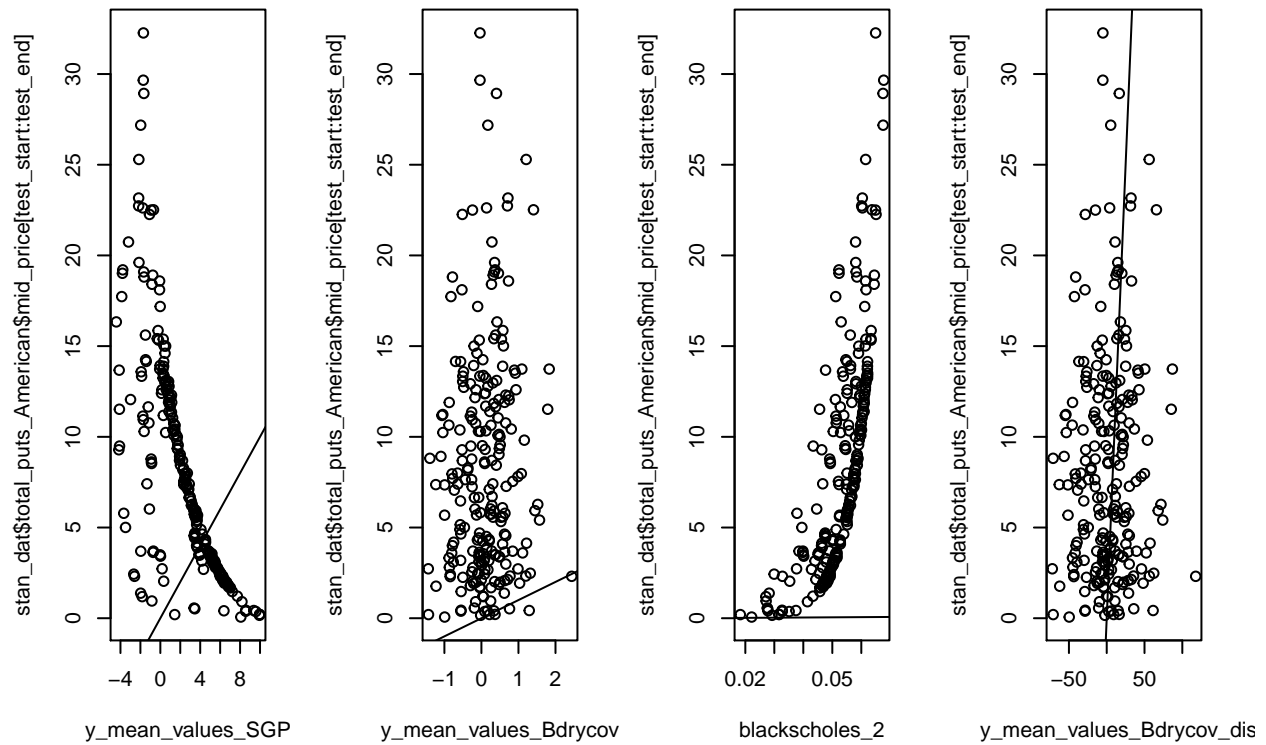
3-3: Plotting Predicted Values against Truth

```
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] 115.4352

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

## [1] 106.3851

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] 992.5983
```