Performance comparison of approximate inference techniques used by WinBugs & Vibes against exact analytical approaches

Anirban Sinha*

Tuesday December 13th 2005

* anirbans@cs.ubc.ca

WinBugs & Vibes

- WinBugs
  - Uses Gibbs sampling.
  - Runs only on Windows.
  - Allows you do draw Bayesian networks (Doodles).
- Vibes
  - Uses Variational mean field sampling.
  - Built using Java, so runs on any platform.
  - Also allows drawing of Bayesian Networks.

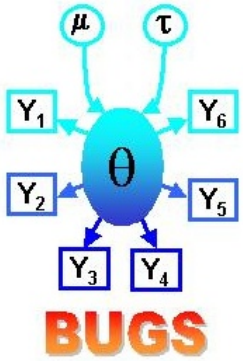
Problem Used to Test the tools:

- Linear Regression
- The same dataset we used in Homework 7.1
  - House prices in Boston area available from the UCI machine learning repository http://www.ics.uci.edu/~mlearn/databases/housing/
  - 506 input data, each data item has 14 columns.
  - I have used the 14th column (house price) as the value to be predicted, and the first 1-13 columns as input features of every data item.

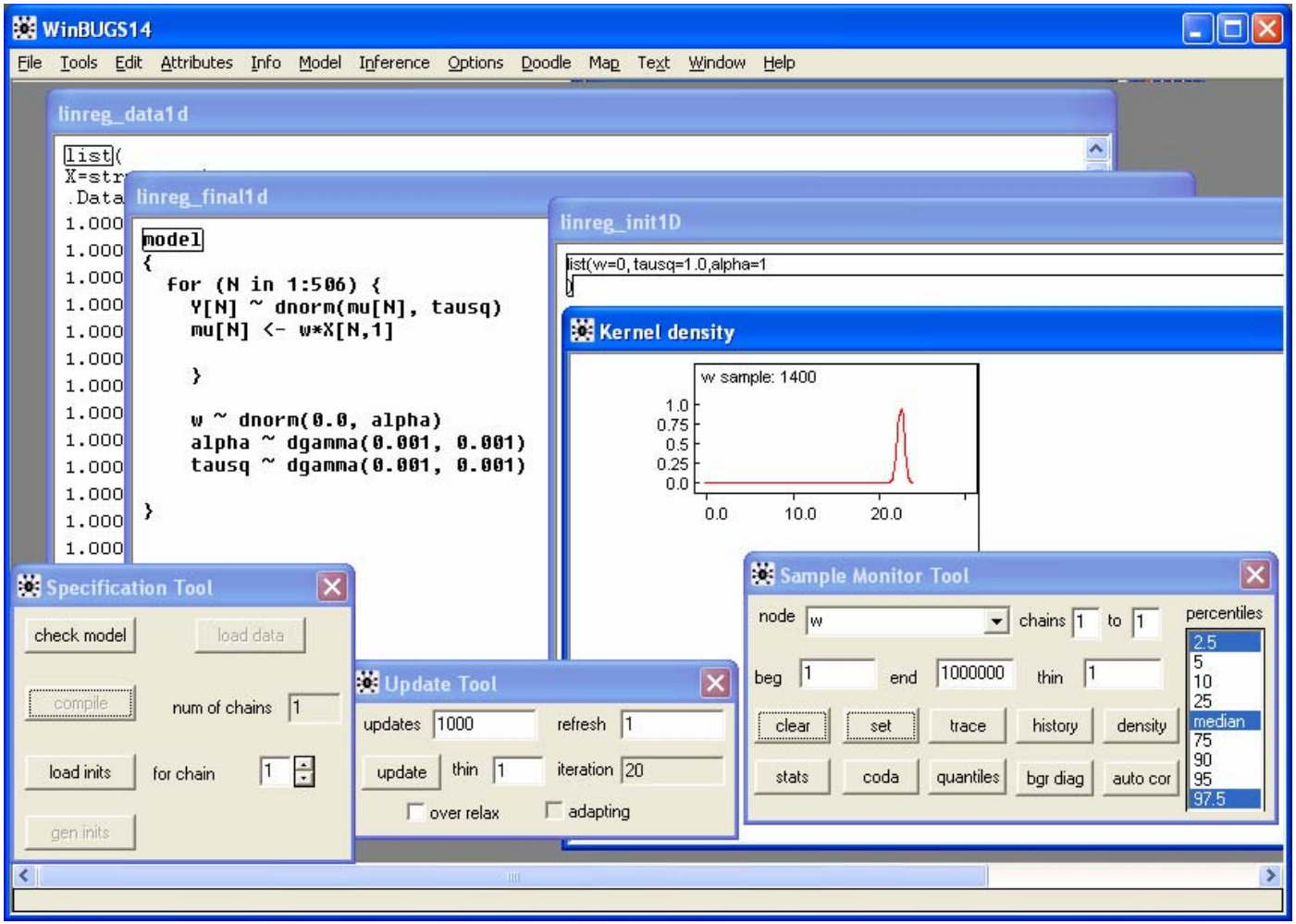
Regression Equation: $$f(x; w) = w_0 + \sum_{i=1}^{13} w_i x_i$$

Model Assumptions & Initializations

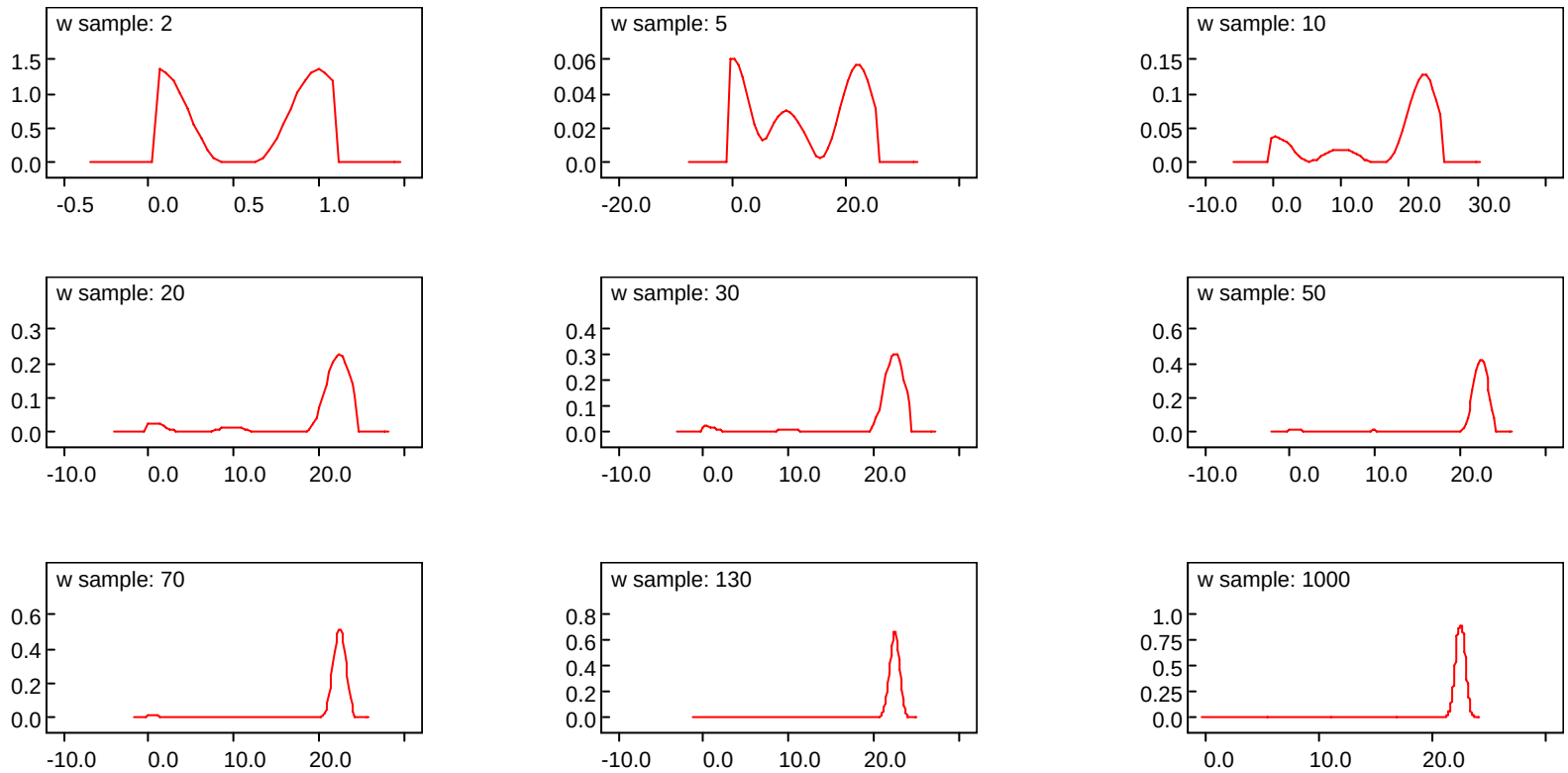
- Weight vector follows Normal Distributions.
- The initial mean is 0.
- For 1-D Gaussian, precision has a gamma prior with a=0.001 & b=0.001
- For 2-D Gaussian, precision has a Wishart prior with R=[1 0; 0 1] & DOF = 2.


μ
τ
Y1
Y6
θ
Y2
Y5
Y3
Y4
BUGS

WinBugs


WinBUGS14
File Tools Edit Attributes Info Model Inference Options Doodle Map Text Window Help
linreg_data1d
linreg_final1d
model
{
for (N in 1:506) {
Y[N] ~ dnorm(mu[N], tausq)
mu[N] <- w*X[N,1]
}
w ~ dnorm(0.0, alpha)
alpha ~ dgamma(0.001, 0.001)
tausq ~ dgamma(0.001, 0.001)
}
linreg_init1D
list(w=0, tausq=1.0,alpha=1
Kernel density
w sample: 1400
1.0
0.75
0.5
0.25
0.0
0.0
10.0
20.0
Specification Tool
check model
load data
compile
num of chains 1
load inits
for chain 1
gen inits
Update Tool
updates 1000
refresh 1
update
thin 1
iteration 20
over relax
adapting
Sample Monitor Tool
node w
chains 1 to 1
percentiles
2.5
5
10
25
median
75
90
95
97.5
beg 1
end 1000000
thin 1
clear
set
trace
history
density
stats
coda
quantiles
bgr diag
auto cor

1-D Linear Regression in WinBugs: Kernel Density Estimate of Posterior



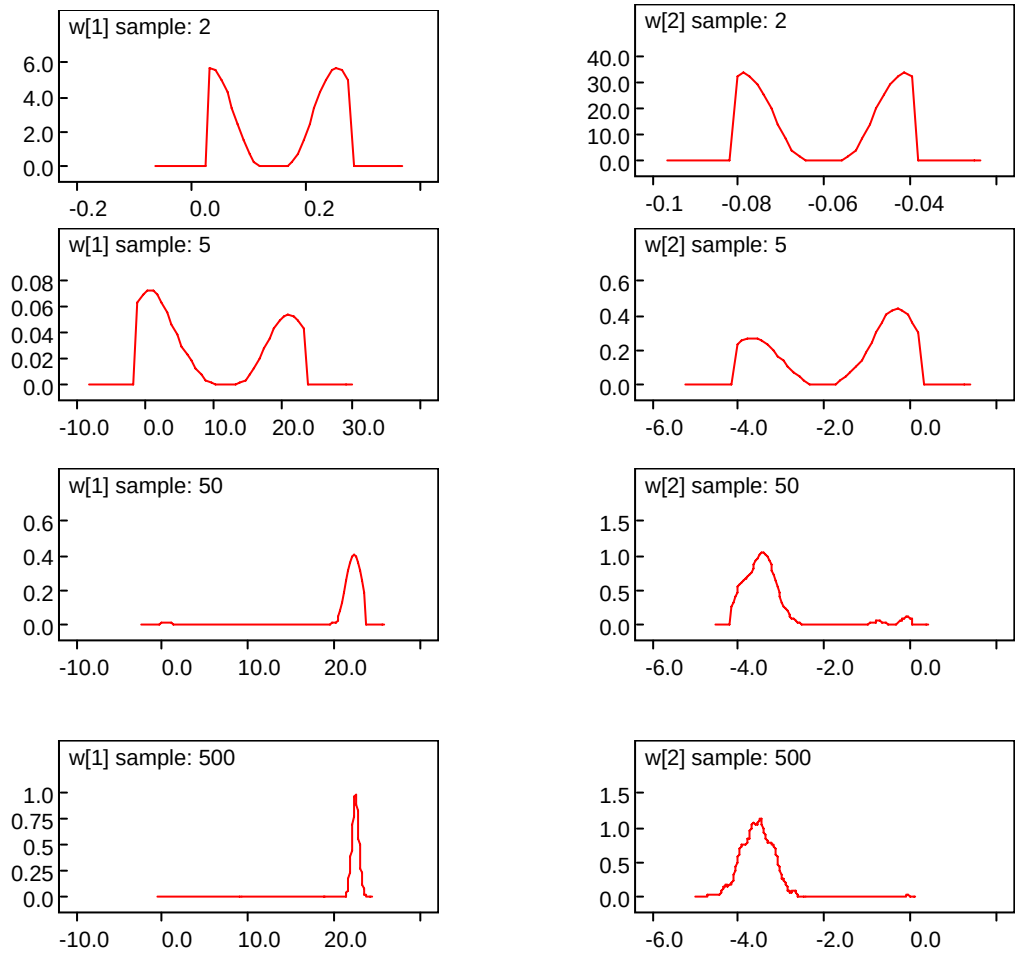
1-D: MAP Estimation Vs. Exact Results

- Final Value Using Bugs:
  - W = 22.1287608458411 (mean of all samples)
- Exact estimation
  - W = 22.5328 (mean of 14^{th} column across all datasets)
- Converges in approximately 100 updates.

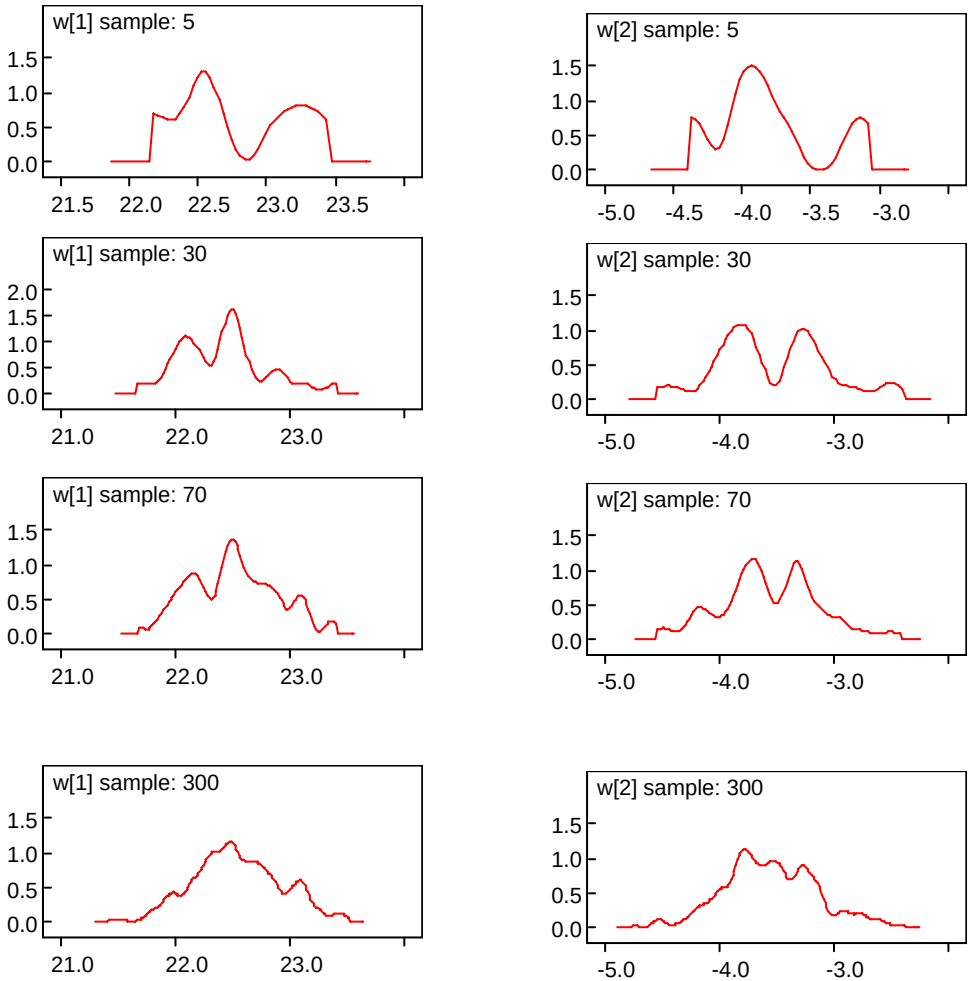
2-D Linear Regression in WinBugs

- Two separate cases analyzed to compare results with Vibes
  - Case 1: Weights are assumed to be uncorrelated (which is not generally the case, but Vibes assumes them to be so).
  - Case 2: The real case where weights are correlated & hence we have a joint Gaussian distribution over all dimensions of w.

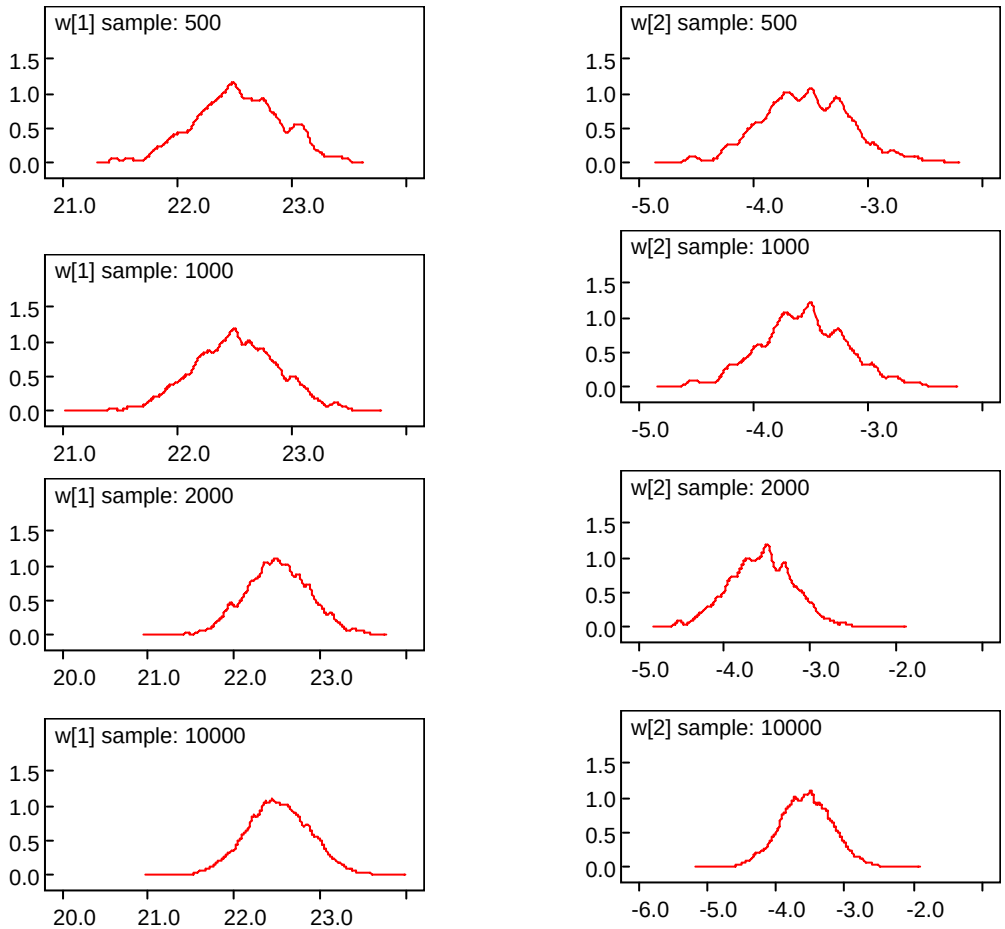
2-D Linear Regression with uncorrelated weights – KDE estimation for each dimension



2-D Linear Regression with correlated weights – KDE estimation of posterior



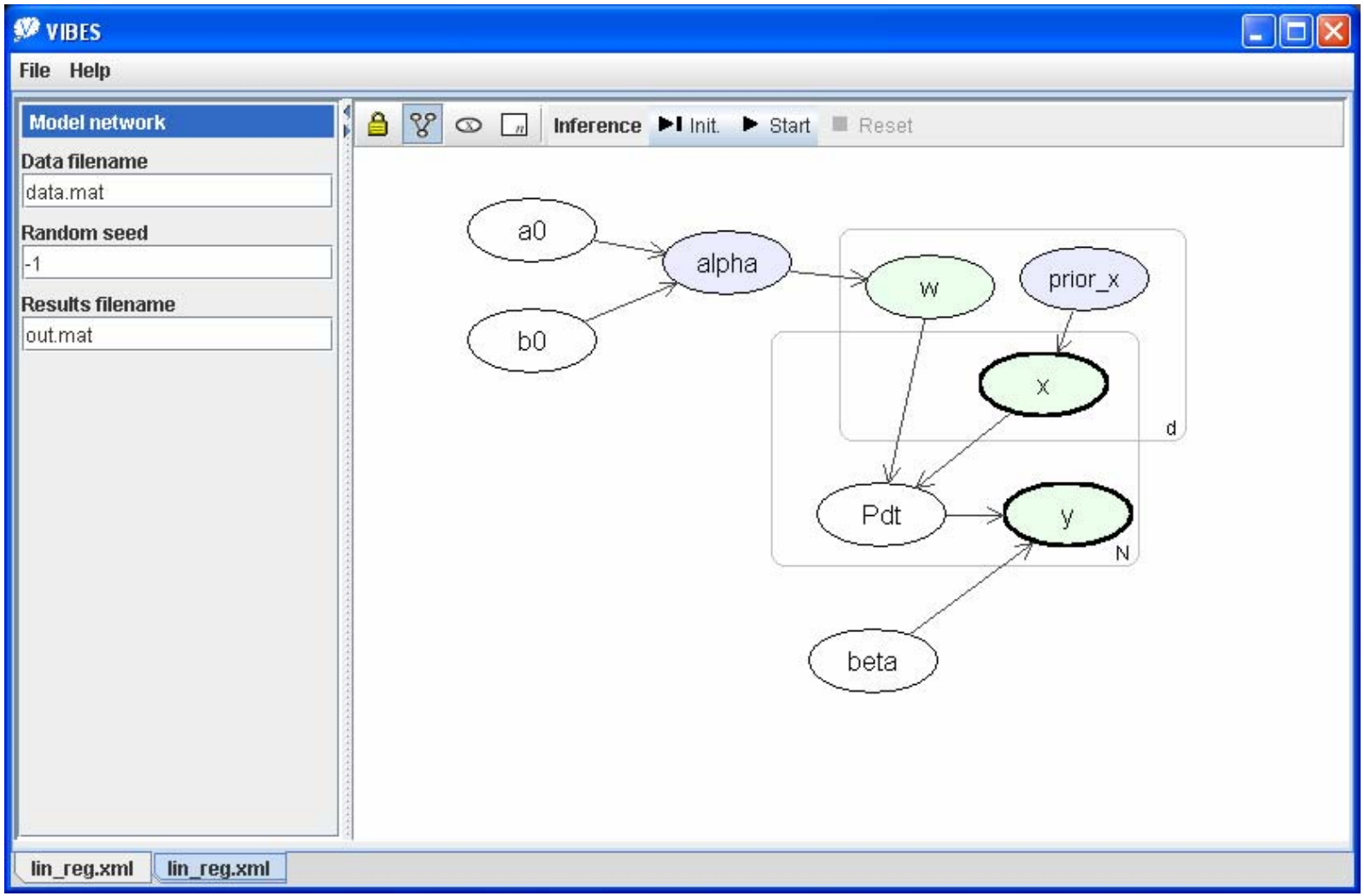
2-D Linear Regression with correlated weights – KDE estimation of posterior – Continued …



2-D: MAP Estimations & Exact Results

- Final MAP estimates using Bugs (uncorrelated)
  - w =[22.14718403715704 -4.07887508183542 9]
  - Converges in 1000 iterations approx.
- Final MAP estimates using Bugs (correlated)
  - w = [22.37869697123022 -2.90123110117053]
  - Converges in 10,000 iterations or more.
- Exact Analytical Results
  - w = [22.309000 -3.357675]

Vibes


VIBES
File Help
Model network
Data filename
data.mat
Random seed
-1
Results filename
out.mat
Inference
Init.
Start
Reset
a0
b0
alpha
w
prior_x
x
d
Pdt
y
N
beta
lin_reg.xml
lin_reg.xml

Vibes Weaknesses

- VIBES doesn't support **conditional density models**, so an uninformative prior on input data is necessary.

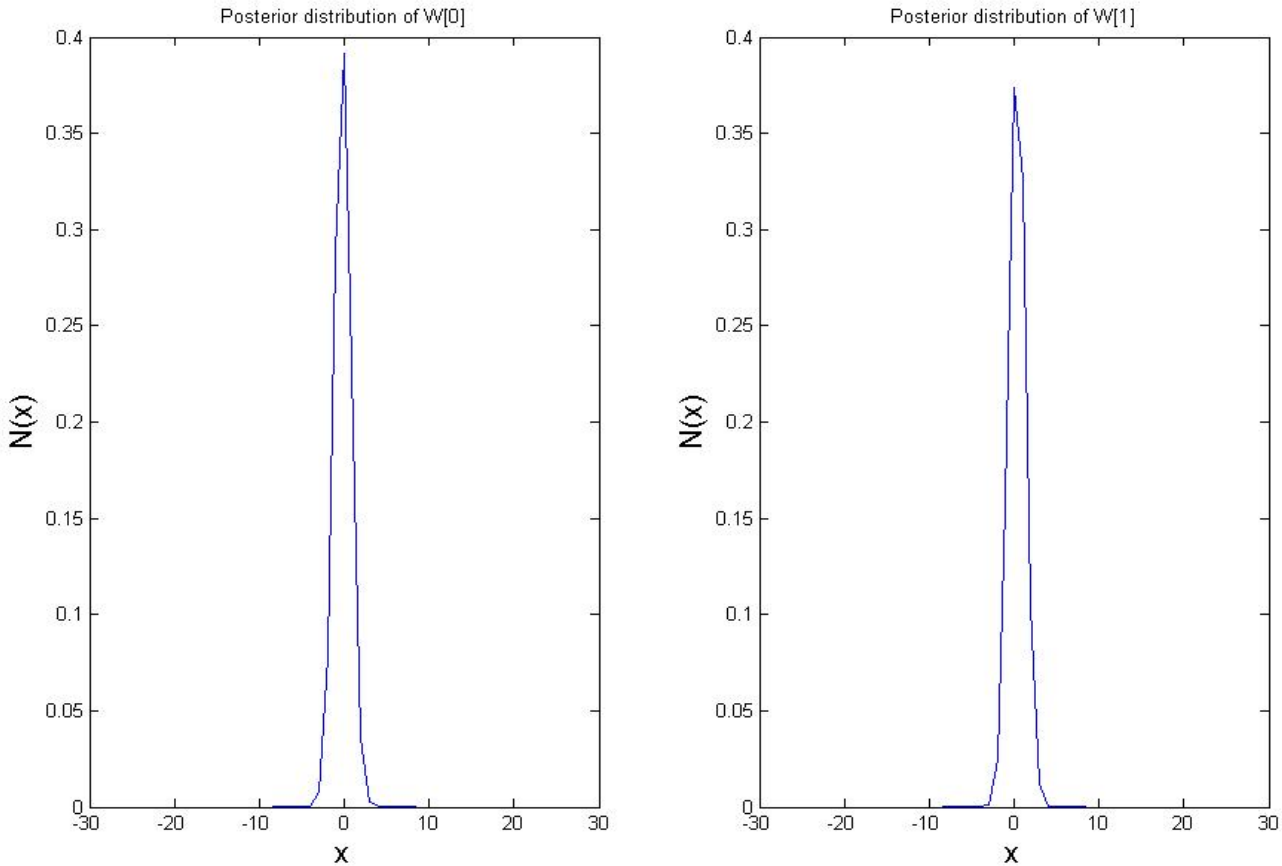
- Vibes does not support multivariate Gaussian posterior distributions. Quoting John Winn in his email to me:

> ***"Sadly, the current version of Vibes does not support multivariate Gaussian posterior distributions. Hence, it is not possible to extract a full covariance matrix. It would be a straightforward project to add multivariate Gaussians to VIBES … Unfortunately, I do not have the time to do this. Apologies …"***

Therefore …

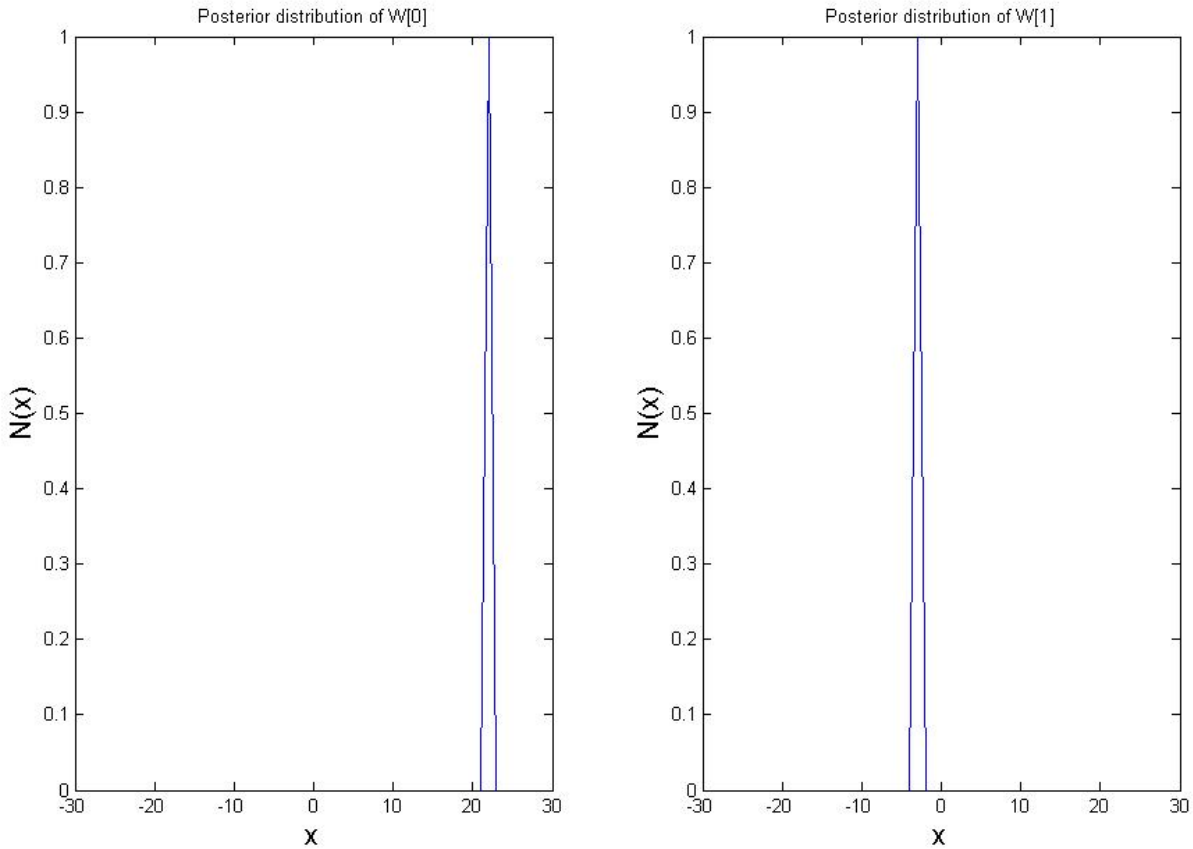
- Our network design is based on 2-D joint posterior distribution.
- However, since we can not extract full covariance matrix, I have taken 1-D plots of the posteriors for each dimensions.
- Also I have taken 2-D plot for the posterior with diagonal covariance matrix.

1-D Gaussian Posterior with each dimension taken separately



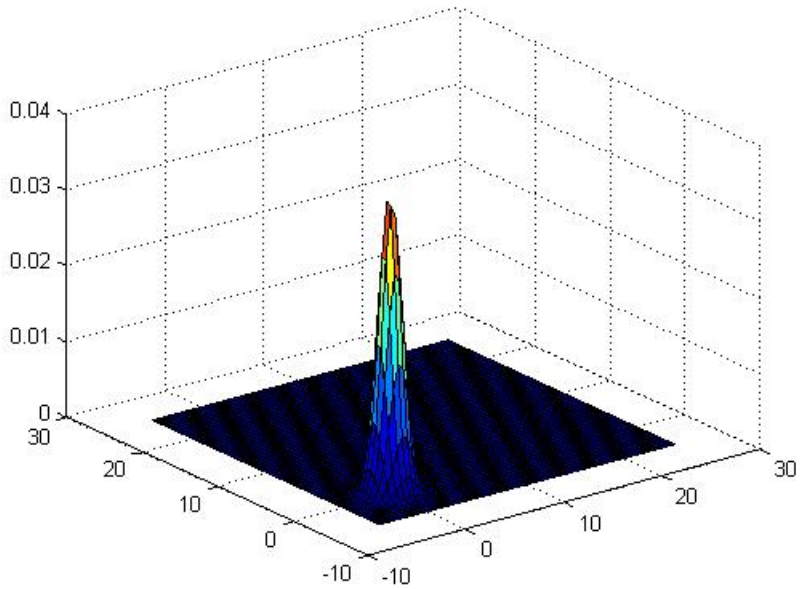
Initialization

1-D Gaussian Posterior with each dimension taken separately

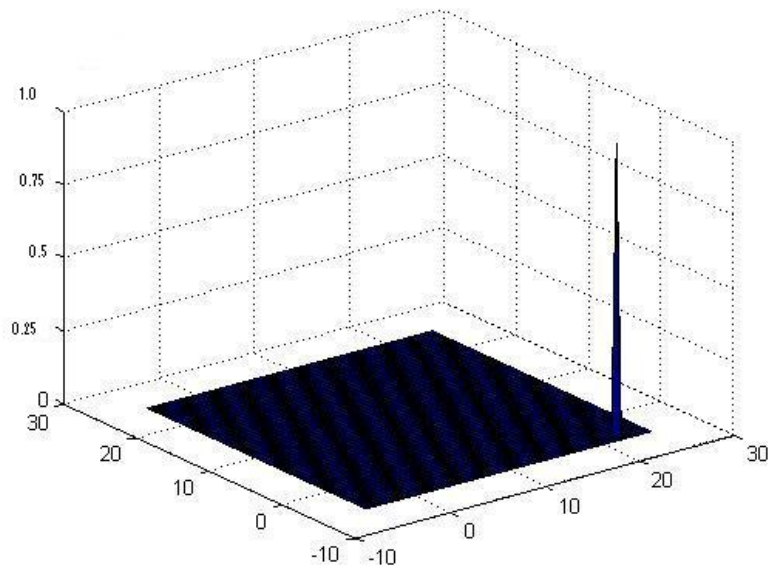


After 1 iteration (converges)

2-D Posterior plots with diagonal covariance matrices



Initialization



After 1 Iteration (converges)

I had also made an AVI demo for it, which did not prove to be very effective because Vibes converges very fast in 2 iterations

MAP Estimations Compared with Exact Estimations

- 2-D weight vectors
  - Estimated w = [22.308965 -3.357670]
  - Exact w = [22.309000 -3.357675]
- 14-D weight vectors
  - Estimated weight vector:

    [22.5317 -0.9289 1.0823 0.1404 0.6825 -2.0580 2.6771 0.0193 -3.1064 2.6630
    -2.0771 -2.0624 0.8501 -3.7470]
  - Exact Weight Vector:

    [22.5328 -0.9291 1.0826 0.1410 0.6824 -2.0588 2.6769 0.0195 -3.1071 2.6649
    -2.0788 2.0626 0.8501 -3.7473]
  - Converges in approximately 88 iterations

Summary

- Vibes performs better in terms of estimated results & number of iterations (speed).
- However it is extremely limited in terms of number of distributions, models supported & available features like plots.
- WinBugs has many diverse features but no direct Matlab interface except if you use MatBugs.
- I did not find ways to plot 3D Gaussians in Bugs. Is there any?

I am grateful to …

- Dr. Kevin Murphy, instructor, CPSC 540
- Frank Hutter, TA, CPSC 540
- John Winn, developer Vibes
- Maryam Mahdaviani
- Alfred Pang

That's all folks

- Codes with instructions to run them are available from my homepage at:

  http://cs.ubc.ca/~anirbans/ml.html

- Questions …