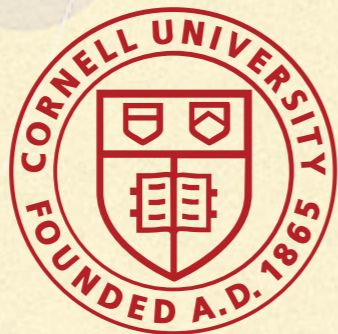
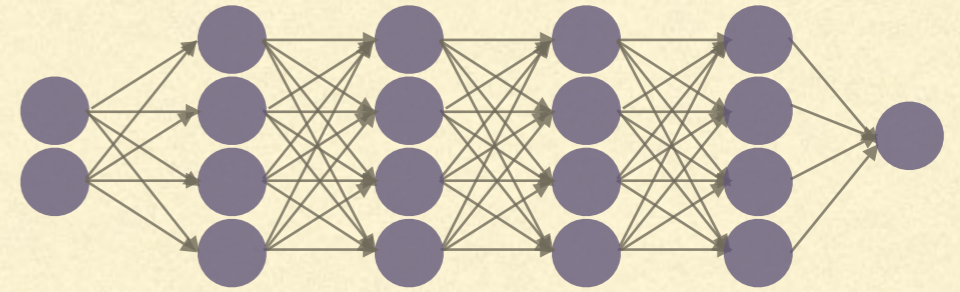

GPYTORCH: BLACKBOX MATRIX- MATRIX GAUSSIAN PROCESS INFERENCE WITH GPU ACCELERATION

Jacob R. Gardner, **Geoff Pleiss**,
David Bindel, Kilian Q. Weinberger, Andrew Gordon Wilson

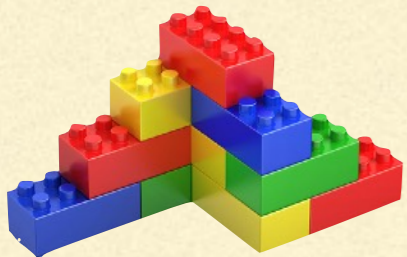


Cornell University

DEEP LEARNING & RAPID PROTOTYPING



1. GPU acceleration

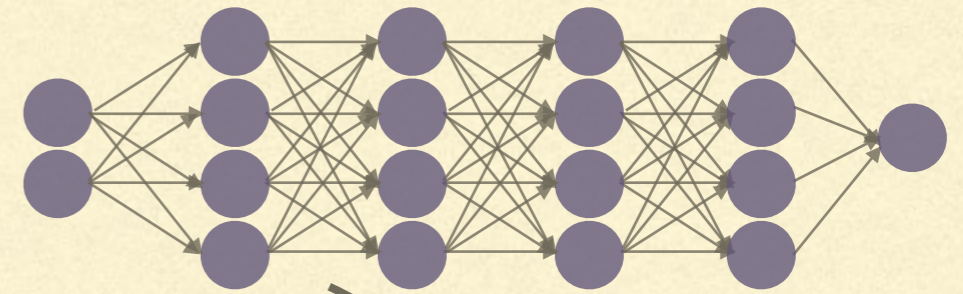


2. Modular design

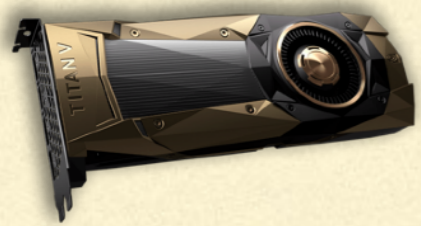


3. Open source software

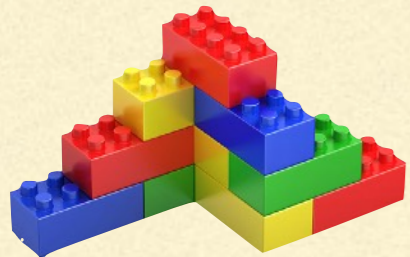
DEEP LEARNING & RAPID PROTOTYPING



$$\left(\begin{matrix} W \end{matrix} \right) \left(\begin{matrix} X \end{matrix} \right)$$



1. GPU acceleration



2. Modular design

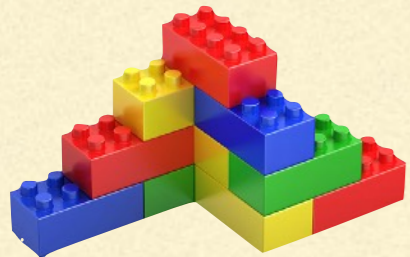


3. Open source software

DEEP LEARNING & RAPID PROTOTYPING



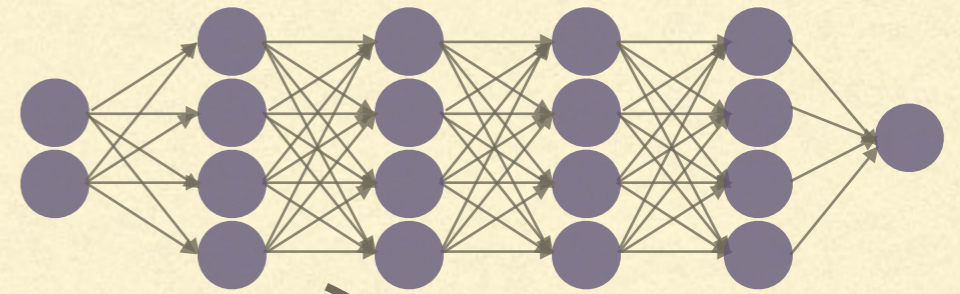
1. GPU acceleration



2. Modular design



3. Open source software



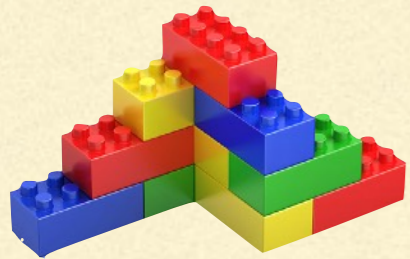
$$\begin{pmatrix} W \end{pmatrix} \begin{pmatrix} X \end{pmatrix}$$

```
model.add_module(  
    nn.Linear(64, 128)  
)
```

DEEP LEARNING & RAPID PROTOTYPING



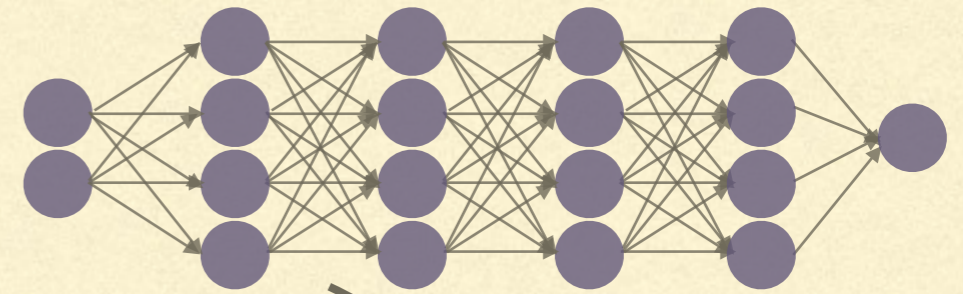
1. GPU acceleration



2. Modular design



3. Open source software

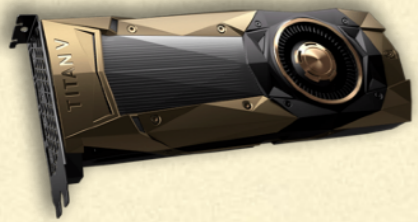
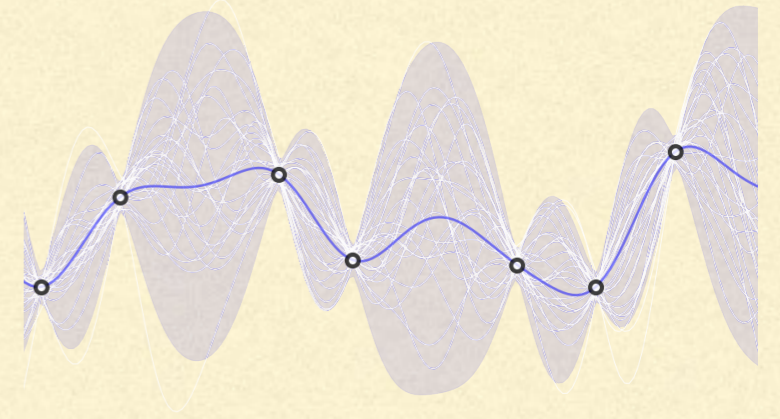


$$(W)(X)$$

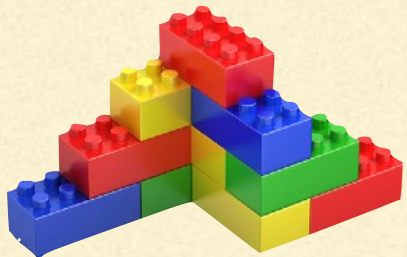
```
model.add_module(  
    nn.Linear(64, 128)  
)
```



GAUSSIAN PROCESSES (GPS) & RAPID PROTOTYPING?

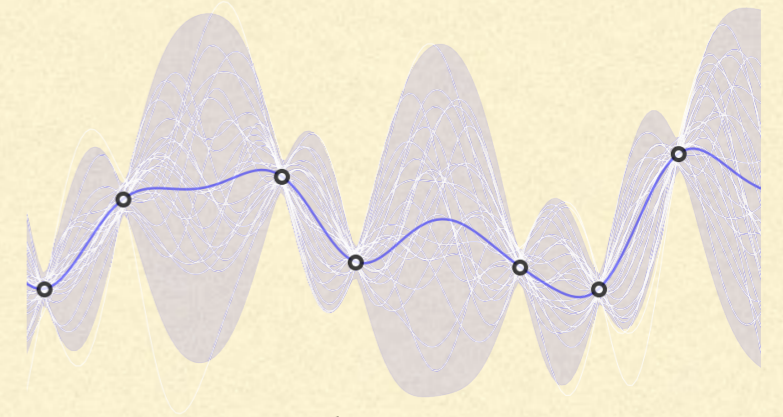


1. GPU acceleration?



2. Modular design?

GAUSSIAN PROCESSES (GPS) & RAPID PROTOTYPING?



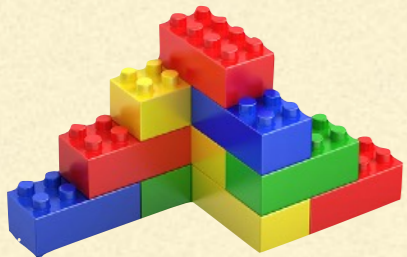
data covariance matrix

$$\log \det(K)$$

$$K^{-1}y$$

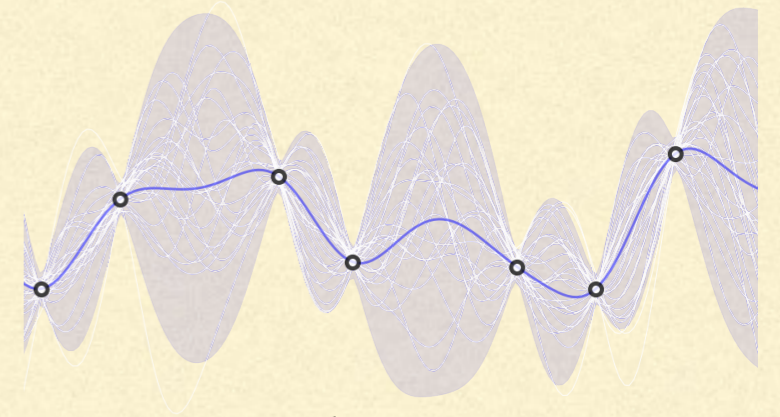


1. GPU acceleration?



2. Modular design?

GAUSSIAN PROCESSES (GPS) & RAPID PROTOTYPING?



data covariance matrix

$$\log \det(K)$$

$$K^{-1}y$$



1. GPU acceleration?



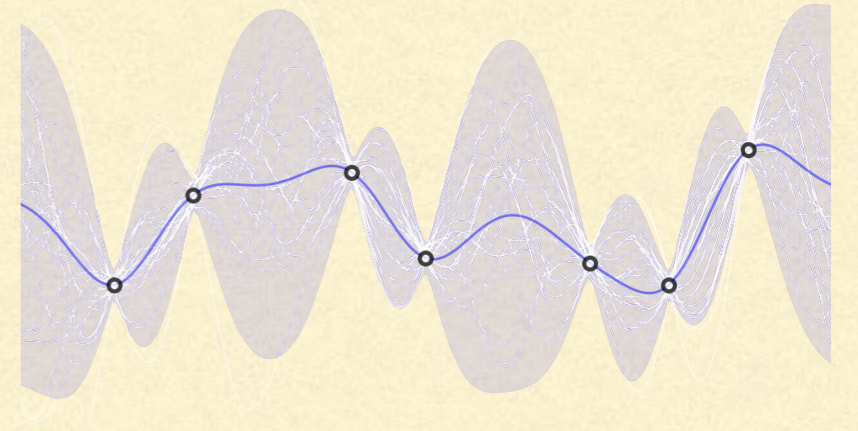
2. Modular design?

$$\frac{1}{(K + \sigma^2 I)^{-1}}$$

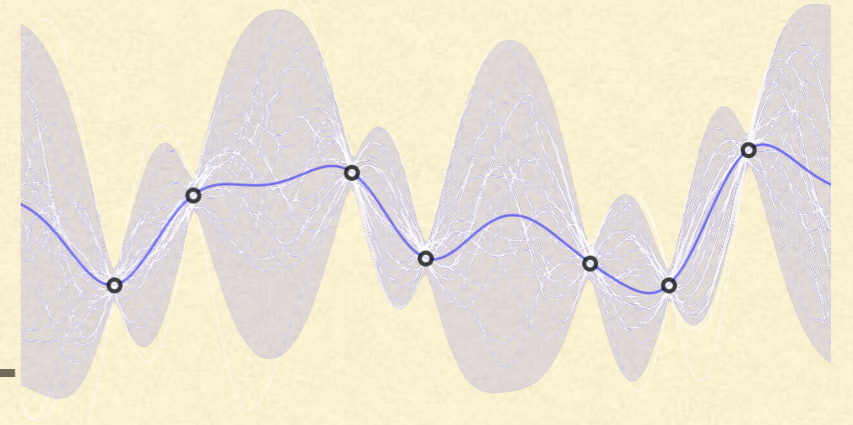


$$\sigma^{-2}I - \sigma^{-4}K_{XU}L^{-\top} (I + L^{-1}K_{XU}K_{XU}L^{-\top}\sigma^{-2})^{-1} L^{-1}K_{UX}$$

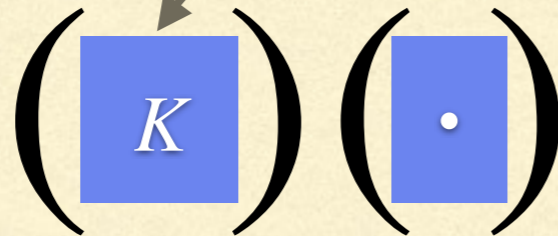
BLACKBOX MATRIX- MATRIX (BBMM) INFERENCE



BLACKBOX MATRIX-MATRIX (BBMM) INFERENCE



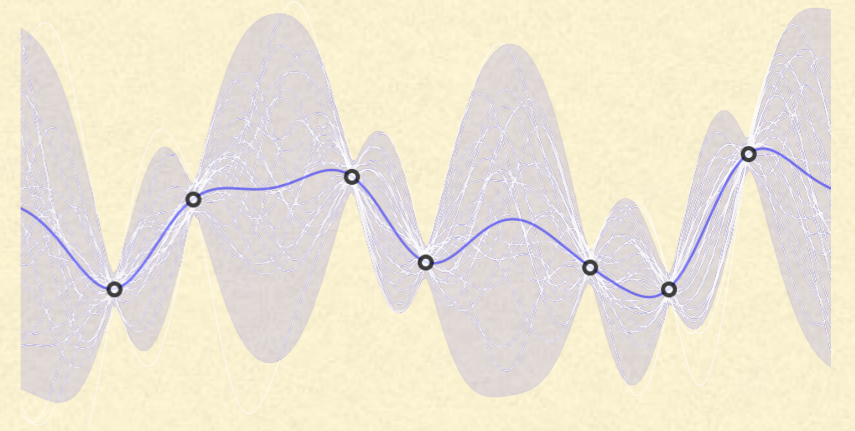
data covariance matrix



(Modified) batch
conjugate gradients

$$\begin{bmatrix} \log \det(K) \\ K^{-1}y \end{bmatrix}$$

BLACKBOX MATRIX-MATRIX (BBMM) INFERENCE

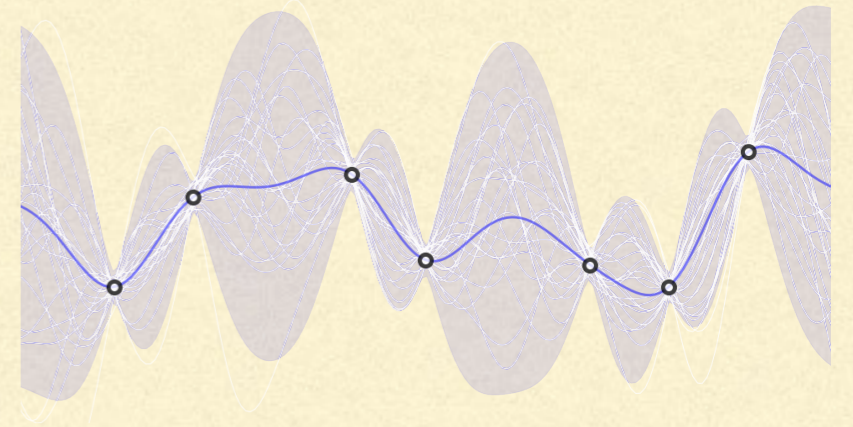


Matrix-matrix
multiplication
w/ K

(Modified) batch
conjugate gradients

$$\begin{bmatrix} \log \det(K) \\ K^{-1}y \end{bmatrix}$$

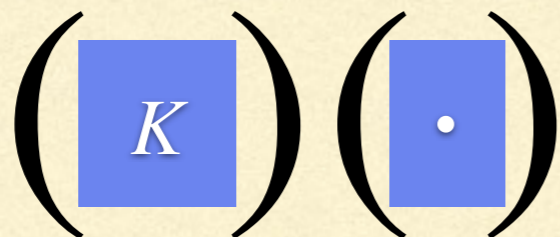
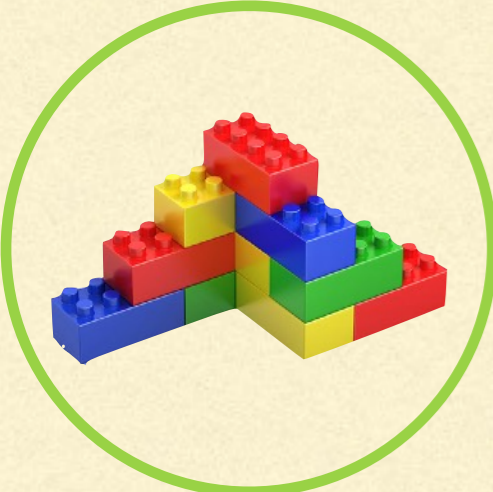
BLACKBOX MATRIX-MATRIX (BBMM) INFERENCE



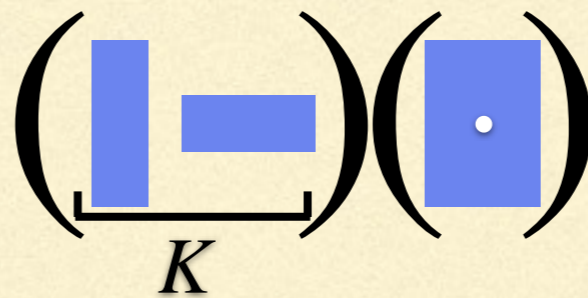
Matrix-matrix multiplication w/ K

(Modified) batch conjugate gradients

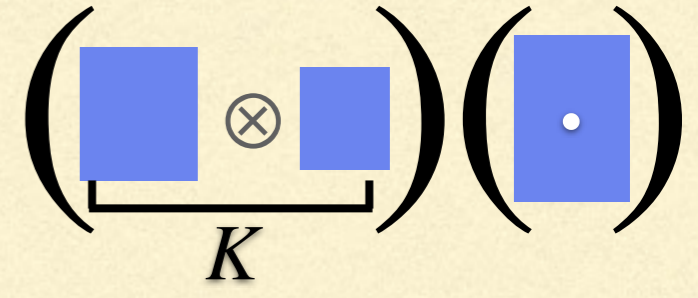
$\begin{bmatrix} \log \det(K) \\ K^{-1}y \end{bmatrix}$



Std. GP



Scalable GP



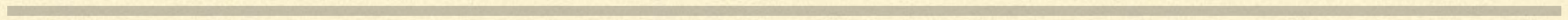
Multi-output GP

$\Theta(n^3)$ $\mathbf{O}(n^2)$

GPYTORCH

A BBMM LIBRARY

GPyTorch



GPYTORCH

A BBMM LIBRARY

GPyTorch



BBMM (Modified CG)

```
In [ ]: train()
```

Standard (Cholesky)

```
In [ ]: train()
```

GPYTORCH

A BBMM LIBRARY

GPyTorch



BBMM (Modified CG)

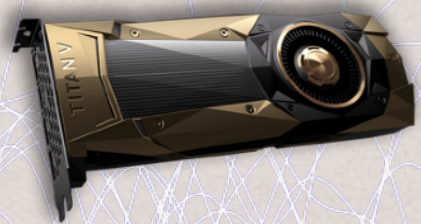
```
In [ ]: train()
```

Standard (Cholesky)

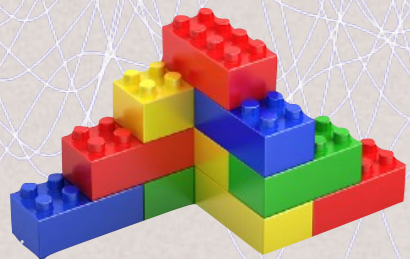
```
In [ ]: train()
```



GPS WITH BBMM INFERENCE



1. GPU acceleration

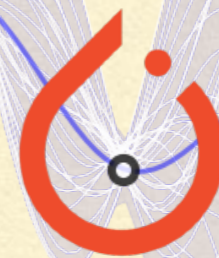


2. Modular design



3. Open source

GPyTorch



Poster #13



<http://gpytorch.ai>