Suppose we have eccess to:  $x_1 \rightarrow y_{11}, -y_{2m}, f(y_{2n}), -y_{3m}$ 27 J 9+1 2 --- > 9 Tm > f (gra), --, f (gra) & one interested in the integrols: (f) p(g|X=x=) ly Sf(y)p(y/X=xr) ly These mend to be estimated z! simultoneously.

Note that:

S & 13) P(g | X=x4) & g  $= \int \langle f, R(.y) \partial_y \rho(y) | X = x^{2} \rangle$  $= \left\langle \int \left\{ \int \left\{ \left( \cdot, y \right) p(y \mid X = x^{\bullet}) \right\} \right\} \right\rangle$ ( ) 4/X= x A reosonable strategy would kherefore be to estimate PY/X= x from samples, Flen use the approximation in the inner product above li. a doina recodrature

is equivalent to approximating a kernel mean).

Note that our data is exactly what is needed for the empirical versions of conditional kernel mean embeddings.

We could ever estimate
the integral for  $x^* \notin \{x_1, ---, x_7\}$ which is pretty wool!

The main difference between the conditional embedding setting & the publication setting is that we observe function values exactly!

 $\lim_{N \to \infty} (x)^{(\cdot)} = k_{\infty} (x_{n} x_{n})^{T} (k_{\infty}(x_{n} x_{n})^{T})$   $+ n \times \lambda_{n}$ 

ky (, 21:m)

 $E_{YIX=x} [f(y)]$   $= \langle f, \hat{f}(y) \rangle$ 

. If we look of WCE over My we only get a difference of conditions? Remed mean enbeddings.

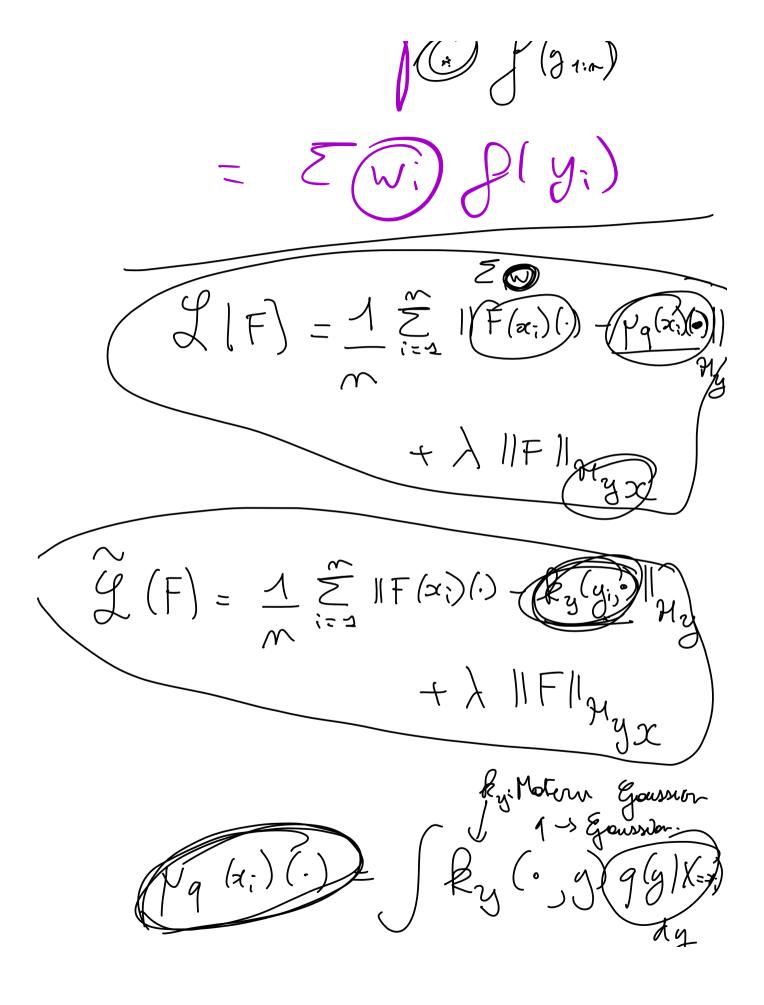
- · Con use Arthur's poper to get rotes in norsy setting?
- Eunctional GP regression.

  with the same pernel of they do vv-regussion.
- o posterior verionce should

  be MCMD evaluated

  et z !

 $= \left( \frac{1}{x} \right)^{T} \left( \frac{1}{$ 



$$\widehat{\mathcal{G}}(F) = \frac{1}{m} \underbrace{\mathbb{E}_{||F(x_i)(\cdot)} \underbrace{\mathbb{E}_{y(y_i,\cdot)}}_{m}}_{\mathcal{H}_y}$$

$$\frac{1}{1180} \left[ \frac{1}{1180} \right] = \frac{1}{1180} \left[ \frac{1}{1180} \left( \frac{1}{1180} \right) \right]$$

$$\frac{1}{1180} \left[ \frac{1}{1180} \left( \frac{1}{1180} \right) \right]$$