Linear reg

$\rightarrow p(\theta) = \mathcal{N}(\theta; \mu_0, \Sigma_0)$

$p(y|X,\theta) = \mathcal{N}(y; \underline{X}\theta, \sigma^2 I_N)$

Maximum likelihood $\theta \leftarrow \arg\max_\theta \log p(y|\theta, X) \rightarrow$ closed form

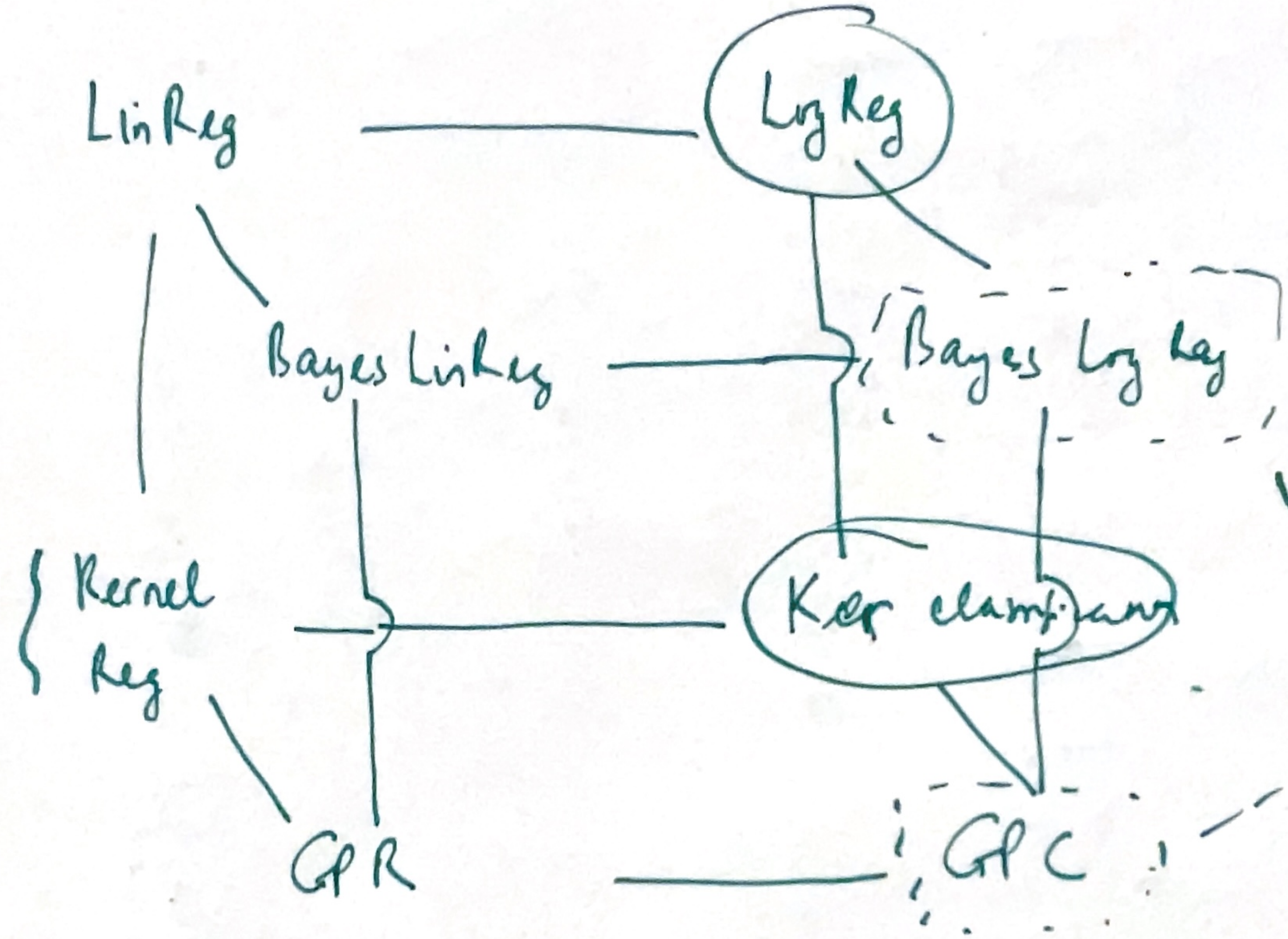
MAP $\theta \leftarrow \arg\max_\theta \log p(y|\theta, X) + \log p(\theta) \leftarrow$ closed form

Bayes exact $p(\theta|y, X) \leftarrow$ closed form

$\log p(\theta|X,y)$

$$H = \frac{d^2 \log p(\theta|X,y)}{d\theta \, . \, d\theta^T} \nearrow D \times D$$

$$H_{ij} = \frac{d^2}{d\theta_i \, d\theta_j} \log p(\theta|X,y)$$




no closed form $p(\theta|X,y)$

$p(f|X,y)$

approx.

- Laplace: $\theta_{MAP}$

  $H|_{\theta=\theta_{map}}$

  $q(\theta) = \mathcal{N}(\theta_{map}, (-H)^{-1})$

- VI $q(\theta) = \mathcal{N}(\theta; \underline{\mu}, \underline{\Sigma})$

  $KL(q(\theta) || p(\theta|X,y))$

  $\downarrow$

  Evidence lower bound

GPR

$$p(f) = GP(m(\cdot), k(\cdot,\cdot)) \quad |\theta$$

$$p(y|f, X, \theta) = \mathcal{N}(y; f(X), \sigma^2 I_N)$$

$$\Rightarrow p(f|y, X) = GP(\hat{m}(\cdot), \hat{k}(\cdot,\cdot))$$

$$p(y|X,\theta) = \mathcal{N}(y; 0, K_{XX} + \sigma^2 I_N) \rightarrow \text{find } \underline{\theta}$$

$M=1$ $\quad B : \underline{D \times 1}$ $\quad b$ $\quad z_n = b^T x_n$ $\quad \tilde{x}_n = b z_n = b \cdot b^T \cancel{z_n}$ $x_n$

$\uparrow$

$\|b\|_2^2 = 1$

Reconstruction $\quad L(z, b) = \frac{1}{N} \sum_n \| \tilde{x}_n - x_n \|_2^2 \leftarrow$ minimise

$$= \frac{1}{N} \sum_n \| b b^T \cancel{z_n} - x_n \|_2^2 \leftarrow$$

$x_n$

data variance matrix

$$= \frac{1}{N} \sum_n (b b^T \cancel{z_n} - x_n)^T (b b^T \cancel{z_n} - x_n)$$

$x \qquad x$

$$= \frac{1}{N} \sum_n \Big[ x_n b \underbrace{b^T \cdot b}_{b^T b = \|b\|_2^2 = 1} b^T x_n - 2 x_n^T b b^T x_n + x_n^T x_n \Big]$$

$b : D \times 1 \qquad b^T x_n : 1 \times 1$

$x_n \; D \times 1 \qquad = x_n^T b$

$$= \frac{1}{N} \sum_n \Big[ x_n^T x_n - \underbrace{x_n^T b}_{} \underbrace{b^T x_n}_{} \Big]$$

$$= \frac{1}{N} \sum_n \Big[ x_n^T x_n - \underbrace{b^T x_n}_{z_n} \underbrace{x_n^T b}_{z_n} \Big] = \frac{1}{N} \sum_n (x_n^T x_n) = \underbrace{b^T \Big( \frac{1}{N} \sum_n x_n \cdot x_n^T \Big) b}_{}$$

$$- \underbrace{b^T \cdot S \, b}_{\text{variance of } z}$$

$$p(y \mid X, \theta) = \prod_n \mathcal{N}(y_n; \theta^T x_n, \sigma^2)$$

Factor analysis $\quad Y \quad N \times D \qquad X\,? : N \times K$

$$p(Y \mid X, \theta) = \prod_n \prod_d \mathcal{N}\Big(y_{n,d}; \underbrace{\boxed{\sum_k \theta_{dk} x_{nk}}}_{\boxed{\theta_d^T \cdot x_n}}, \sigma^2\Big)$$

$$p(X) = \prod_n p(X_n) = \prod_n \underbrace{\mathcal{N}(x_n; 0, I_K)}$$

$\theta$ is known

$$\underline{p(X \mid Y, \theta)}$$

$$\underbrace{\log p(Y \mid \theta)}$$

Probabilistic PCA

$\sigma^2 \to 0$

$X, \theta \to PCA$

GPLVM GP Latent variable model

$$p(Y \mid X, \theta) = \prod_n \prod_d \mathcal{N}(y_{n,d}; f_d(x_n), \sigma^2)$$

$$p(f) = \prod_d GP(f_d; m(\cdot), k(\cdot, \cdot))$$

$$p(X) = \prod_n \mathcal{N}(x_n; 0, I_K)$$