3 Régression linéaire

- Pour le modèle de la *régression linéaire aux moindres carrés*, énoncer les équations pour : l'hypothèse du modèle ; la fonction d'erreur ; la valeur du vecteur des pondérations qui minimise la fonction d'erreur

$$\begin{aligned}
\text{modèle : } t &= \vec{w}^{\mathrm{T}}\vec{\phi}(\vec{x}) \\
\text{fonction d'erreur } : E(\vec{w}) &= \frac{1}{2}\sum_{n=1}^{N}\left(t_n - \vec{w}^{\mathrm{T}}\vec{\phi}(\vec{x}_n)\right)^2 \\
\text{solution : } \vec{w} &= (\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\vec{\mathsf{t}} \\
\text{où } \Phi_n &= \vec{\phi}(\vec{x}_n)^{\mathrm{T}}
\end{aligned}$$

- Pour le modèle de la *régression linéaire (probabiliste)*, énoncer les équations pour : les hypothèses du modèle ; la fonction d'erreur ; la valeur du vecteur des pondérations qui minimise la fonction d'erreur

$$\begin{aligned}
\text{modèle : } y(\vec{x}) &= \vec{w}^{\mathrm{T}}\vec{\phi}(\vec{x}) \\
p(t\,|\vec{x}) &= \mathcal{N}(t\,|y(\vec{x}), \beta^{-1}) \\
\text{fonction d'erreur } : E(\vec{w}) &= \frac{1}{2}\sum_{n=1}^{N}\left(t_n - \vec{w}^{\mathrm{T}}\vec{\phi}(\vec{x}_n)\right)^2 \\
\text{solution : } \vec{w} &= (\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^{\mathrm{T}}\vec{\mathsf{t}} \\
\text{où } \Phi_n &= \vec{\phi}(\vec{x}_n)^{\mathrm{T}}
\end{aligned}$$

- Pour le modèle de la *régression linéaire bayésienne*, énoncer les équations pour : les hypothèses du modèle ; la distribution *a posteriori* du vecteur des pondérations ; la distribution prédictive de la variable cible dans le cas où $\boldsymbol{S}_0 = \alpha\boldsymbol{I}$ et $\vec{m}_0 = \vec{0}$

$$\begin{aligned}
\text{modèle : } y(\vec{x}, \vec{w}) &= \vec{w}^{\mathrm{T}}\vec{\phi}(\vec{x}) \\
p(\vec{w}) &= \mathcal{N}(\vec{w}\,|\vec{m}_0, \boldsymbol{S}_0) \\
p(t\,|\vec{x}, \vec{w}) &= \mathcal{N}(t\,|y(\vec{x}, \vec{w}), \beta^{-1}) \\
\text{distribution } \textit{a posteri } : p(\vec{w}\,|\vec{\mathsf{t}}, \mathbf{X}) &= \mathcal{N}(\vec{w}\,|\vec{m}_N, \boldsymbol{S}_N) \\
\text{où } \boldsymbol{S}_N^{-1} &= \boldsymbol{S}_0^{-1} + \beta\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi} \\
\vec{m}_N &= \boldsymbol{S}_N(\boldsymbol{S}_0^{-1}\vec{m}_0 + \beta\boldsymbol{\Phi}^{\mathrm{T}}\vec{\mathsf{t}}) \\
\text{prédictive : } p(t\,|\vec{x}, \vec{\mathsf{t}}, \mathbf{X}) &= \mathcal{N}(t\,|\vec{m}_N^T\vec{\phi}(\vec{x}), \sigma_N^2(\vec{x})) \\
\text{où } \sigma_N^2(\vec{x}) &= \frac{1}{\beta} + \vec{\phi}(\vec{x})^{\mathrm{T}}\boldsymbol{S}_N\vec{\phi}(\vec{x})
\end{aligned}$$

- Établir la correspondance entre le paramètre de régularisation dans la fonction d'erreur des moindres