问题：

假设数据集为 $\mathcal{D} = \{\mathcal{X}_i, \mathcal{Y}_i\}_{i=1}^{n}$,其中$\mathcal{X}_i \in \mathbb{R}^d, d > 1; \mathcal{Y}_i \in \mathbb{R}$ ，且参数$\mathcal{W}$的先验分布为$\mathcal{W} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\alpha}^{-1}\mathrm{I})$ ，似然函数为 $\mathrm{P}(\mathcal{D}|\mathcal{W}) \triangleq \mathrm{P}(\mathcal{Y}|\mathbb{X}, \mathcal{W}) = \mathcal{N}(\mathcal{Y}|\mathbb{X}\mathrm{w},\ \sigma^2\mathrm{I})$,

证明：MAP(极大后验估计)等同于ℓ_2正则化(岭回归)

提示：1、$\mathrm{P}(\mathcal{W}|\mathcal{D}) = \frac{\mathrm{P}(\mathcal{D}|\mathcal{W})\mathrm{P}(\mathcal{W})}{\mathrm{P}(\mathcal{D})} \propto \mathrm{P}(\mathcal{D}|\mathcal{W})\mathrm{P}(\mathcal{W})$

2、多元高斯概率密度函数：

$$f(x|\mu, \Sigma) = (2\pi)^{-\frac{d}{2}}(\det\Sigma)^{-\frac{1}{2}} \exp\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\}$$

解答：

由题目，$\mathcal{W} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\alpha}^{-1}\mathrm{I_d})$，所以

$$\mathrm{P}(\mathcal{W}) = (2\pi)^{-\frac{d}{2}}(\det\boldsymbol{\alpha}^{-1}\mathrm{I})^{-\frac{1}{2}}\exp\{-\frac{\alpha}{2}\mathrm{w}^T\mathrm{w}\};$$

又 $\mathrm{P}(\mathcal{D}|\mathrm{w}) \triangleq \mathrm{P}(\mathcal{Y}|\mathbb{X}, \mathcal{W}) = \mathcal{N}(\mathcal{Y}|\mathbb{X}\mathrm{w},\ \sigma^2\mathrm{I})$,

故 $P(\mathcal{D}|w)$

$$=(2\pi)^{-\frac{n}{2}}(\det\sigma^2 I_n)^{-\frac{1}{2}}\exp\{-\frac{1}{2\sigma^2}(\mathcal{Y}-\mathbb{X}w)^T(\mathcal{Y}-\mathbb{X}w)\};$$

所以 MAP$\triangleq \underset{w}{\operatorname{argmax}} P(w|\mathcal{D}) \triangleq \underset{w}{\operatorname{argmax}} \frac{P(\mathcal{D}|w)P(w)}{P(\mathcal{D})}$

$$\propto \underset{w}{\operatorname{argmax}} P(\mathcal{D}|w)P(w)$$

$$=\underset{w}{\operatorname{argmax}} \frac{\exp\{-\frac{1}{2\sigma^2}(\mathcal{Y}-\mathbb{X}w)^T(\mathcal{Y}-\mathbb{X}w)-\frac{\alpha}{2}w^Tw\}}{(2\pi)^{\frac{n+d}{2}}(\det\boldsymbol{\alpha}^{-1}I)^{\frac{1}{2}}(\det\sigma^2 I_n)^{\frac{1}{2}}}$$

由于上式分母中不含 w，且 $f(x) = e^x$为单调递增函数，所以

$$\text{MAP}\triangleq \underset{w}{\operatorname{argmax}}\{-\frac{1}{2\sigma^2}(\mathcal{Y}-\mathbb{X}w)^T(\mathcal{Y}-\mathbb{X}w)-\frac{\alpha}{2}w^Tw\}$$

$$\triangleq \underset{w}{\operatorname{argmin}}\{\frac{1}{2\sigma^2}(\mathcal{Y}-\mathbb{X}w)^T(\mathcal{Y}-\mathbb{X}w)+\frac{\alpha}{2}w^Tw\}$$

$$\triangleq \underset{w}{\operatorname{argmin}}\{(\mathcal{Y}-\mathbb{X}w)^T(\mathcal{Y}-\mathbb{X}w)+\alpha\sigma^2 w^Tw\},$$

该式为岭回归形式，式中$\alpha\sigma^2$对应岭回归中超参数λ，因而 MAP 等同于岭回归，原命题得证。