Homework # 3 Solution

Problem 1

$$\begin{aligned}
A^T(AA^T+\lambda I_n)^{-1} &- (A^TA+\lambda I_m)^{-1}A^T \\
&= (A^TA+\lambda I_m)^{-1}\left((A^TA+\lambda I_m)A^T - A^T(AA^T+\lambda I_n)\right)(AA^T+\lambda I_n)^{-1} \\
&= (A^TA+\lambda I_m)^{-1}\left((A^TAA^T+A^T) - (A^TAA^T+A^T)\right)(AA^T+\lambda I_n)^{-1} \\
&= (A^TA+\lambda I_m)^{-1}(0)(AA^T+\lambda I_n)^{-1} = 0
\end{aligned}$$

Problem 2

All norms ($p \geq 1$) satisfy the triangle inequality, i.e. $\|x+y\|_p \leq \|x\|_p + \|y\|_p$. Hence, if we take a convex combination, the convexity follows directly from the triangle inequality.

$$\|\alpha x + (1-\alpha)y\|_p \leq \|\alpha x\|_p + \|(1-\alpha)y\|_p = \alpha\|x\|_p + (1-\alpha)\|y\|_p, \forall \alpha \in [0,1].$$

If x and y have entries such that the sign of x_i is the same as the sign of y_i, then $|x_i + y_i| = |x_i| + |y_i|$. Hence for such sign aligned vectors

$$\|\alpha x + (1-\alpha)y\|_1 = \|\alpha x\|_1 + \|(1-\alpha)y\|_1 = \alpha\|x\|_1 + (1-\alpha)\|y\|_1, \forall \alpha \in [0,1].$$

Hence the 1-norm is not strictly convex.

Problem 3

$p(b|x)$ is is $N(Ax, R_n)$ and $p(x)$ is is $N(0, R_x)$. The MAP estimate is given by

$$x_{MAP} = \arg\max_x \left[\log p(b|x) + \log p(x)\right].$$

Substituting for the distributions and dropping terms that do not contribute to the optimization process, we have

$$x_{MAP} = \arg\min_x \left[(b-Ax)^T R_n^{-1}(b-Ax) + x^T R_x^{-1} x\right]$$

This can be readily solved by setting the derivative with respect to x to zero.

$$x_{MAP} = (A^T R_n^{-1} A + R_x^{-1})^{-1} A^T R_n^{-1} b.$$