Problem 9.1

a) Let $v, w \in \mathcal{M}$ and $t \in [0,1]$. f is convex, therefore

$$f(tv + (1-t)w) \le \underbrace{t f(v) + (1-t) f(w)}_{= tm + (1-t)m = m}$$

because v and w are minimizers of f.

By definition of m we have $m \le f(tv + (1-t)w)$

hence $f(tv + (1-t)w) = m$: $tv + (1-t)w \in \mathcal{M}$.

Conclusion: $\mathcal{M}$ is convex.

b) By contradiction, assume that there exists two distincts minimizers $v \neq w$.

By strict convexity of f:

$$f\left(\frac{v+w}{2}\right) < \frac{1}{2} f(v) + \frac{1}{2} f(w) = m$$

which contradicts the definition of m.

Conclusion: f has a unique minimizer

Problem 9.2

a) Let $x \in \mathbb{R}^n$,
$$\begin{cases} \nabla f(x) = 2Mx + b \\ Hf(x) = 2M. \end{cases}$$

f is convex if and only if $Hf(x)$ is positive semi definite for all x, that is, if and only if M is positive semi definite.

b) f is convex and differentiable, therefore:

f admits a minimizer $\iff$ there exists $x \in \mathbb{R}^n$ such that $\nabla f(x) = 0$.

$\iff$ $2Mx + b = 0$

$\iff$ there exists $u \in \mathbb{R}^n$ such that $Mu = b$

$\iff$ $b \in \mathrm{Im}(M)$.

Problem 9.3

a) The function $h : x \mapsto \|x\|^2$ is strictly convex since for all $x \in \mathbb{R}^n$, $H_h(x) = 2 I d_n$ is positive definite.

Let f be a strongly convex function. Let $\alpha > 0$ and g convex such that

$$f(x) = g(x) + \frac{\alpha}{2}\|x\|^2 \text{ for all } x \in \mathbb{R}^n$$

Let $x, y \in \mathbb{R}^n$ and $t \in [0,1]$. By convexity of g and strict convexity of h we have:

- $g(tx + (1-t)y) \leq t g(x) + (1-t) g(y)$
- $\|tx + (1-t)y\|^2 \leq t\|x\|^2 + (1-t)\|y\|^2$

with a strict inequality when $t \in (0,1)$ and $x \neq y$.

Combining the inequalities, we get that $f(tx + (1-t)y) \leq t f(x) + (1-t) f(y)$ with strict inequality whenever $t \in (0,1)$ and $x \neq y$.

Conclusion: f is strictly convex.

b) For $\alpha > 0$ we define $f_\alpha(x) = \varphi(x) - \frac{\alpha}{2}\|x\|^2$

Notice that f_α is twice differentiable, with Hessian given by:

$$H_{f_\alpha}(x) = H_\varphi(x) - \alpha \, Id_n$$

φ is strongly convex

$\iff$ there exists $\alpha > 0$ such that f_α convex

$\iff$ there exists $\alpha > 0$ such that for all $x \in \mathbb{R}^n$ the eigenvalues of $H_{f_\alpha}(x)$ are non-negative.

$\iff$ there exists $\alpha > 0$ such that for all $x \in \mathbb{R}^n$ the eigenvalues of $H_\varphi(x)$ are all $\geqslant \alpha$.

The last equivalence comes from the fact that the eigenvalues of $H_{\varphi(x)} - \alpha \, Id_n$ are the eigenvalues of $H_\varphi(x)$ minus α.

Problem 9.4

a) $f(x) = \|Ax-y\|^2 = (Ax-y)^T(Ax-y)$

$$= x^T A^T A x - 2y^T A x + \|y\|^2$$

From 9.2, we know that $\begin{cases} \nabla f(x) = 2A^T A x - 2y^T A y \\ Hf(x) = 2A^T A \end{cases}$

Since we know that $2A^TA$ is positive semidefinite (since $v^T A^T A v = \|Av\|^2 \geq 0$ for all $v \in \mathbb{R}^n$) we get that f is convex.

b) By the rank - nullity theorem:

$$\dim \mathrm{Ker}(A) = m - \mathrm{rank}(A) \geq 1 \text{ if } \mathrm{rank}(A) < m$$

In that case, one can find a non-zero vector v in $\mathrm{Ker}(A)$.

Note that
- $f(v) = \|Av - y\|^2 = \|y\|^2 = f(0)$.
- similarly, $f((1-t)v) = f(0)$.

Therefore, for $t \in [0,1]$,

$$f(t\cdot 0 + (1-t)v) = f(0) = t f(0) + (1-t) f(v)$$

Since $v \neq 0$, f is not strictly convex.

c) If $\mathrm{rank}(A) = m$, then the first m singular values of A are all strictly positive: $\sigma_1 \geq \dots \geq \sigma_m > 0$.

A^TA is $m \times m$: its eigenvalues are then $\sigma_1^2 \geq \dots \geq \sigma_m^2 > 0$.

Since for all $x \in \mathbb{R}^m$, the eigenvalues of $Hf(x) = 2A^TA$ are $\geq 2\sigma_m^2$ which is > 0, we get that f is strongly convex.

Problem 9.5

Let us fix $x \in \mathbb{R}^n$.

We define the function

$$g : \mathbb{R}^n \longrightarrow \mathbb{R}$$
$$\underline{\underline{h}} \longmapsto f(x+h) - f(x) - \nabla f(x) \cdot h - \frac{\gamma}{2} \|h\|^2$$

- g is twice differentiable, and its Hessian is:

$$H_g(h) = H_f(x+h) - \gamma \, Id_n$$

Since the eigenvalues of $H_f(x+h)$ are $\geqslant \gamma$ (by definition of γ) we get that $H_g(h)$ is positive semi definite for all h: g is convex.

- $\nabla g(h) = \nabla f(x+h) - \nabla f(x) - \gamma h$

hence $\nabla g(0) = 0$: since g is convex, 0 is a minimizer of g. Note that $g(0) = 0$, hence for all $h \in \mathbb{R}^n$, $g(h) \geqslant 0$:

$$f(x+h) \geqslant f(x) + \nabla f(x) \cdot h + \frac{\gamma}{2} \|h\|^2.$$

The 2nd inequality follows from the same arguments