PROBLEM 9.1

$\mathcal{M} = \{ v \in \mathbb{R}^n, f(v) = m \}$

(a) $\mathcal{M}$ is a convex set:

Let v_1 and v_2 belong to $\mathcal{M}$. By definition:

$$f(v_1) = f(v_2) = m$$

$\forall \alpha \in [0,1]$, $f(\alpha v_1 + (1-\alpha) v_2) \leqslant \alpha f(v_1) + (1-\alpha) f(v_2)$

$\Rightarrow \forall \alpha \in [0,1] \quad f(\alpha v_1 + (1-\alpha) v_2) \leqslant m$

yet m is a minimizer: $m \leqslant f(\alpha v_1 + (1-\alpha) v_2)$

$\Rightarrow f(\alpha v_1 + (1-\alpha) v_2) = m \quad \Rightarrow \quad \alpha v_1 + (1-\alpha) v_2 \in \mathcal{M}$ -

$\mathcal{M}$ is a convex set -

(b) Assume v_1 and v_2 are different minimizers of f:

If f is strictly convex

$$\alpha f(v_1) + (1-\alpha) f(v_2) > f(\alpha v_1 + (1-\alpha) v_2)$$

$\Rightarrow \quad m > m \quad$ contradiction

PROBLEM 9.2

$$f(x) = x^T M x + b^T x + c$$

(a) $f(x) = \sum_{ij} x_i M_{ij} x_j + \sum_i b_i x_i + c$

$$\nabla f(x) = \begin{pmatrix} \vdots \\ \sum_j 2 M_{ij} x_j + b_i \\ \vdots \end{pmatrix} \leftarrow \text{coordinate } i \quad = 2Mx + b$$

$$[H_f(x)]_{ii} = 2M_{ii}$$

$$[H_f(x)]_{ij} = 2M_{ij}$$

$\Rightarrow \quad H_f(x) = 2M.$

$\downarrow$

f is convex if and only if M is psd.

(b) Since we assume M psd, f is a convex function.

We saw in class that for convex functions:

x is a minimizer $\iff \nabla f(x) = 0$

Hence if x is a minimizer $\Rightarrow 2Mx + b = 0$

$\Rightarrow b = M(-2x) \Rightarrow b \in \mathrm{Im}(M).$

And if $b \in \mathrm{Im}(M)$: there exist $x \in \mathbb{R}^n$ such that $b = Mx$

$\Rightarrow M(-x) = -b.$

$\Rightarrow \nabla f(-2x) = 0$

$\Rightarrow f$ admits a minimizer.

PROBLEM 9.3

(a) $f : x \longmapsto \|x\|^2$ is strictly convex :

$$\nabla f(x) = \begin{vmatrix} 2x_1 \\ \vdots \\ 2x_n \end{vmatrix} \qquad H_f(x) = 2\,\mathrm{Id}_n \longrightarrow \text{positive definite}$$

$\Rightarrow x \longmapsto \|x\|^2$ is strictly convex.

(+) Remark that for $f(x) = g(x) + h(x)$, twice differentiable functions,

$$\boxed{H_f(x) = H_g(x) + H_h(x).}$$

So $H_f(x) = H_g(x) + 2\,\mathrm{Id}_n$.

If $H_g(x)$ psd, then all its eigenvalues $\lambda_1 \geq, \lambda_2 \cdots \lambda_n \geq 0$.

$\Rightarrow$ Eigenvalues of $H_f(x)$ which are $\lambda_1 + 2 \geq, \lambda_2 + 2 \geq, \cdots \lambda_n + 2 > 0$ are positive

$\Rightarrow H_f(x)$ is positive definite.

$\Rightarrow f$ is strictly convex.

(b) If $\alpha > 0$ exists: $\quad \varphi(x) = \underbrace{\varphi(x) - \alpha\|x\|^2}_{g(x)} + \alpha\|x\|^2$

by calculations similar as above,
$H_g(x)$ psd as non negative e.v.

$\Rightarrow g(x)$ convex $\Rightarrow \varphi$ strongly convex.

If φ is strongly convex: $\quad f(x) = g(x) + \alpha\|x\|^2$

$\hookrightarrow$ by computations above H_f eigenvalues $> \alpha$.

PROBLEM 9.4.

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

↳ use the results of 9.2 (a). $\left|\begin{array}{l} \nabla f(x) = 2A^TAx - 2A^Ty \\ H_f(x) = 2A^TA \end{array}\right.$

(b) $\dim \ker(A) \geq 1$ $\Rightarrow$ we have a non zero v such that $Av = 0$

$\Rightarrow f(0) = \|y\|^2 = f(v) = f(tv)$ for $t \in [0,1]$

so $f(tv + (1-t)0) = f(0) = tf(0) + (1-t)f(v)$

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

(c) if $\operatorname{rank}(A) = m$, $f(x) = \|Ax - y\|^2$

$H_f(x) = 2A^TA$ → eigenvalues of A^TA, squared singular values of A:

$\sigma_1^2 \geq \sigma_2^2 \cdots \sigma_m^2 > 0$

→ PROBLEM 9.3 (b) with $\alpha = 2\sigma_m^2$.

PROBLEM 9.5

Let $f(x) = \ln(e^{x_1} + e^{x_2} + \cdots + e^{x_n})$.

Compute $H_f(x)$ and conclude whether f is convex.

$$f_2(x) = \ln(e^{x_1} + \cdots + e^{x_n})$$

$$\nabla f_2(x) = \frac{1}{e^{x_1} + \cdots e^{x_n}} \nabla (e^{x_1} + \cdots e^{x_n})$$

$$= \frac{1}{e^{x_1} + \cdots + e^{x_n}} \begin{vmatrix} e^{x_1} \\ \vdots \\ e^{x_n} \end{vmatrix} = \begin{vmatrix} \partial f_2/\partial x_1 \\ \vdots \\ \partial f/\partial x \end{vmatrix}$$

$$\partial f_{ii} = -\frac{1}{(e^{x_1} + \cdots + e^{x_n})^2} (e^{x_i})^2 + \frac{e^{x_i}}{e^{x_1} + \cdots + e^{x_n}}$$

$$= \frac{}{(e^{x_1} + \cdots + e^{x_n})^2} \left(e^{x_i} (e^{x_1} + \cdots + e^{x_n}) - (e^{x_i})^2 \right)$$

$$= \frac{\sum_{j \neq i} e^{x_j + x_i}}{\left(\sum_j e^{x_j} \right)^2}$$

$$\partial f_{ij} = -\frac{1}{(\quad)^2} e^{x_i} e^{x_j}$$

$$\beta(x) = e^{x_1} + \cdots + e^{x_n}$$

$$H = \frac{1}{\beta^2} \begin{pmatrix} \sum_{j \neq i} e^{x_j + x_i} & & -e^{x_i + x_j} \\ & \ddots & \end{pmatrix}$$

$$\begin{pmatrix} \ddots & & \\ & \ddots & \\ -e^{x_i+x_j} & & \ddots \end{pmatrix}$$

$$z^T H z = \frac{1}{\beta^2} z^T \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$$

$$= \frac{1}{\beta^2} z^T \begin{pmatrix} \sum_{j \neq 1} e^{x_j+x_1} z_1 - \sum_{j\neq 1} e^{x_j+x_1} z_j \\ \vdots \end{pmatrix}$$

$$= \frac{1}{\beta^2} z^T \begin{pmatrix} \sum_{j \neq 1} e^{x_j+x_1} (z_1 - z_j) \\ \vdots \end{pmatrix}$$

$$= \frac{1}{\beta^2} \sum_i \sum_j e^{x_j+x_i} (z_i - z_j) z_i$$

$$= \frac{1}{\beta^2} \sum_i \sum_j e^{x_j+x_i} (z_j - z_i) z_j$$

$$= -\frac{1}{\beta^2} \sum_i \sum_j e^{x_j+x_i} (z_i - z_j) z_j .$$

$$= \frac{1}{} \sum \sum e^{x_j+x_i} (z_i - z_j)^2 . \geq 0 .$$

$2\beta^2$ $\quad i \quad j$

so that the function is convex.