COMP/ENGN 8535

Homework 4 Solutions

Problem 1

By definition we have that

$$\mathrm{kurt}(X+Y) = E[(X+Y)^4] - 3\left(E[(X+Y)^2]\right)^2 \quad ①$$

Consider the first term on RHS of ①:

$$E[(X+Y)^4]$$

binomial theorem/ first-principles expansion

$$= E[X^4 + 4X^3Y + 6X^2Y^2 + 4XY + Y^4]$$

linearity of expectation

$$= E[X^4] + 4E[X^3Y] + 6E[X^2Y^2] + 4E[XY] + E[Y^4]$$

X & Y indp.

$$= E[X^4] + 4E[X^3]\underbrace{E[Y]}_{0} + 6\underbrace{E[X^2]}_{1}\underbrace{E[Y^2]}_{1} + 4\underbrace{E[X]}_{0}\underbrace{E[Y]}_{0} + E[Y^4]$$

Given $E[X] = E[Y] = 0$ and $E[X^2] = E[Y^2] = 1$.

$$= E[X^4] + E[Y^4] + 6 \quad ②$$

Similarly the second term on the RHS of ① is

$$3\left(E\left[(X+Y)^2\right]\right)^2 = 3\left(E\left[X^2 + 2XY + Y^2\right]\right)^2$$

expansion

$$= 3\left(E[X^2] + E[2XY] + E[Y^2]\right)^2$$

X & Y indp.

$$= 3\left(E[X^2] + 2E[X]E[Y] + E[Y^2]\right)^2$$

$$= 3(1 + 0 + 1)^2$$

$$= 3 \cdot 2^2$$

$$= 12. \quad ③$$

∴ Combining both ② & ③ on RHS of ① gives

$$\text{kurt}(X+Y)$$

$$= E[X^4] + E[Y^4] + 6 - 12$$

$$= E[X^4] + E[Y^4] - 6$$

noting $E[X^2] = E[Y^2] = 1$
so $3(E[X^2])^2 + 3(E[Y^2])^2 = 6$.

$$= E[X^4] - 3\left(E[X^2]\right)^2 + E[Y^4] - 3\left(E[Y^2]\right)^2$$

$$= \text{kurt}(X) + \text{kurt}(Y)$$

$\boxed{\text{Problem 2}}$

a) $$I(Y_1;Y_2) = \iint p(y_1,y_2) \log \frac{p(y_1,y_2)}{p(y_1)p(y_2)} \, dy_1 \, dy_2$$

$$= \iint p(y_1,y_2) \log p(y_1,y_2) \, dy_1 \, dy_2$$

$\log \frac{a}{bc} = \log a - \log b - \log c$

$$- \iint p(y_1,y_2) \log p(y_1) \, dy_1 \, dy_2$$

$$- \iint p(y_1,y_2) \log p(y_2) \, dy_1 \, dy_2$$

$$= -H(Y_1,Y_2) - \int p(y_1) \log p(y_1) \, dy_1$$

sum rule & defn. of joint entropy.

$$- \int p(y_2) \log p(y_2) \, dy_2$$

defn. of entropy

$$= -H(Y_1,Y_2) + H(Y_1) + H(Y_2) \quad (6)$$

b) By definition of negentropy,

$$J(Y_1) = H(Y_g) - H(Y_1) \Rightarrow H(Y_1) = H(Y_g) - J(Y_1) \quad (4)$$

$$J(Y_2) = H(Y_g) - H(Y_2) \Rightarrow H(Y_2) = H(Y_g) - J(Y_2) \quad (5)$$

$\therefore$ (4) & (5) $\rightarrow$ (6) implies

$$I(Y_1;Y_2) = -H(Y_1,Y_2) + 2H(Y_g) - J(Y_1) - J(Y_2)$$

$$= C - \sum_{i=1}^{2} J(Y_i)$$

with $C = 2H(Y_g) - H(Y_1,Y_2)$

Problem 3

a) Let $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

The constraint $\Phi X = y$ with $\Phi = [4, 1]$, $y = 1$ implies that

$$4x_1 + x_2 = 1 \Rightarrow x_2 = 1 - 4x_1$$

so the constrained optimisation problem

$\min_X \|X\|_2$ s.t. $\Phi X = y$ reduces to the unconstrained optimisation

$$\min_{x_1} \sqrt{x_1^2 + (1-4x_1)^2}$$

which shares the same minimising arguments as

$$\min_{x_1} \left[x_1^2 + 1 - 8x_1 + 16x_1^2 \right]$$

$$= \min_{x_1} \left[17x_1^2 - 8x_1 + 1 \right]$$

$f(x_1) = 17x_1^2 - 8x_1 + 1$ is convex function so first order optimality condition $\frac{df}{dx_1} = 0$ is sufficient.

$\therefore$ Optimising x_1^* satisfies

$$\frac{df}{dx_1}(x_1^*) = 0 = 34x_1^* - 8$$

$$\Rightarrow x_1^* = \frac{4}{17}.$$

Thus $x_2^* = 1 - 4x_1^* = 1 - \frac{16}{17} = \frac{1}{17}$.

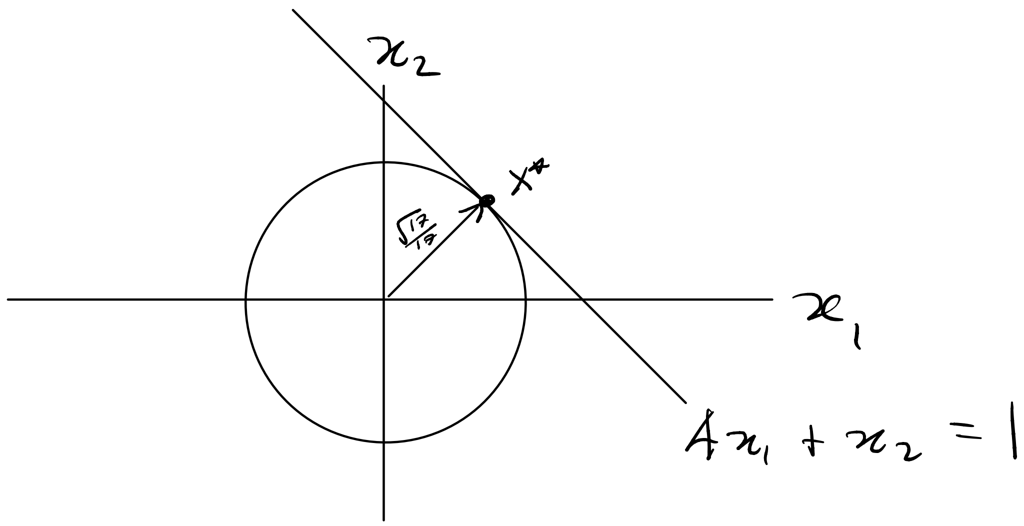
Optimal X^* for original constrained optimisation problem is

$$X^* = \begin{bmatrix} 4/17 \\ 1/17 \end{bmatrix}$$

with optimal value

$$\max_{X} \|X\|_2 \quad \text{s.t.} \quad \Phi X = y$$

$$= \|X^*\| = \frac{\sqrt{17}}{17} = \frac{1}{\sqrt{17}} = \sqrt{\frac{1}{17}}.$$



Computer plots ok for HW but not exam.

b) $\min_{x \in \mathbb{R}^2} \|x\|_1$ s.t. $\begin{bmatrix} 4 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 1$ (7)

Solution

The constraint implies that the only feasible $x \in \mathbb{R}^2$ must satisfy

$$4x_1 + x_2 = 1$$

$$\Rightarrow x_2 = 1 - 4x_1$$

thus the constrained optimisation problem (7) reduces to the unconstrained problem

$$\min_{x_1 \in \mathbb{R}} \left[|x_1| + |1 - 4x_1| \right].$$

Consider subsets of the domain $x_1 \in \mathbb{R}$ over which the absolute value simplifies:

(1) For $x_1 \leq 0$, note that

$$|x_1| = -x_1$$

$$|1 - 4x_1| = 1 - 4x_1$$

$$\therefore \min_{x_1 \le 0} \left[|x_1| + |1-4x_1| \right] = \min_{x_1 \le 0} \left[1 - 5x_1 \right]$$

$$= 1 \text{ at } x_1 = 0.$$

(2) For $0 \le x_1 \le \frac{1}{4}$, note that

$$|x_1| = x_1$$

$$|1-4x_1| = 1-4x_1$$

$$\therefore \min_{0 \le x_1 \le \frac{1}{4}} \left[|x_1| + |1-4x_1| \right] = \min_{0 \le x_1 \le \frac{1}{4}} \left[1 - 3x_1 \right]$$

$$= 1 \text{ at } x_1 = 0.$$

(3) For $x_1 \ge \frac{1}{4}$, note that

$$|x_1| = x_1$$

$$|1-4x_1| = 4x_1 - 1$$

$$\therefore \min_{x_1 \ge \frac{1}{4}} \left[|x_1| + |1-4x_1| \right] = \min_{x_1 \ge \frac{1}{4}} \left[5x_1 - 1 \right]$$

$$= \frac{1}{4} \text{ at } x_1 = \frac{1}{4}.$$

Now to combine the cases note that trivially

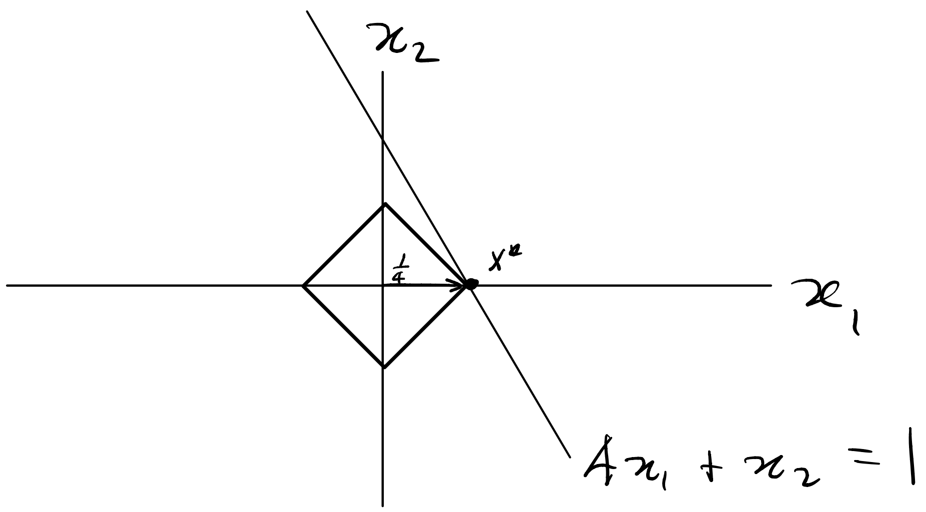
$$\min_{x_1 \in \mathbb{R}} \left[|x_1| + |1-4x_1| \right]$$

$$= \min \left\{ 1,\ 1,\ \tfrac{1}{4} \right\}$$

$= \dfrac{1}{4}$ which is achieved at $x_1 = \dfrac{1}{4}$.

$\therefore$ optimal $x_2 = 1 - 4x_1 = 0$.

so optimal $x = \begin{bmatrix} \frac{1}{4} \\ 0 \end{bmatrix}$, with optimal value $\|x^*\|_1 = \frac{1}{4}$.



Problem 4

a) spark $(\Phi) = 3$ because the 3 columns

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$ are lineary dependent

with $$\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} - \sqrt{2}\begin{bmatrix} 1 \\ 0 \end{bmatrix} - \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

or similarly because the 3 columns

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$ are linearly dependent

with $$\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} - \sqrt{2}\begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

<u>And</u> because no set of 1 or 2 columns are linearly dependent.

(Note that in this case we also have that spark $(\Phi) =$ rank $(\Phi) + 1 = 3$ but this is not always the case so computing rank(Φ) then adding 1 to get the spark(Φ) is not a suitable approach.)

Using the Theorem for uniqueness of sparse solutions to $\Phi x = b$ we have that since spark$(\Phi) = 3 > 2$ then

$\Phi x = b$ will have a unique 1-sparse solution.

Alternative method is to show/recognise that every subset of 2 columns of Φ are linearly independent then use Tao's result stated without spark.

b) Let $b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ then

$\Phi X = b$

$$\Rightarrow \begin{bmatrix} 1 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Note then that a 2-sparse solution $x \in \mathbb{R}^4$ with support $\Omega = \{i, j\}$ for any $1 \leq i, j \leq 4$ defines a system of linear equations

$$\Phi_\Omega x_\Omega = b$$

where $\Phi_\Omega \in \mathbb{R}^{2\times 2}$ is a matrix with columns of Φ corresponding to Ω and $x_\Omega \in \mathbb{R}^2$ is vector with non-zero elements of $x \in \mathbb{R}^4$. That is,

$$\Phi_\Omega = \begin{bmatrix} \Phi_i & \Phi_j \end{bmatrix} \quad \text{and} \quad x_\Omega = \begin{bmatrix} x_i \\ x_j \end{bmatrix}$$

with $\Omega = \{i, j\}$.

Because every pair of columns of Φ are independent (spark$(\Phi) = 3$) the matrix Φ_Ω will be non-singular (rank$(\Phi_\Omega) = 2$) for any 2-sparse support $\Omega = \{i, j\}$.

Thus standard linear algebra results imply that the linear system

$$\Phi_\Omega x_\Omega = b$$

has at most one solution for any 2-sparse support $\Omega = \{i, j\}$. (could have no solution if equations inconsistent).

The possible 2-sparse supports $\Omega = \{i, j\}$ for $1 \le i, j \le 4$

are $\{1,2\}, \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}, \{3,4\}$ so 6 systems Also ${}^4C_2$.

$\Phi_\Omega x_\Omega = b$ with at most one solution each.

Thus, $\Phi x = b$ has at most 6 2-sparse solutions.

Problem 5

a) $$\Phi x = \begin{bmatrix} 1 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 3 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 3 \end{bmatrix} = b$$ verifying $x = \begin{bmatrix} 0 \\ 0 \\ 3 \\ 0 \end{bmatrix}$ solves $\Phi x = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$.

b) $\Phi x = b$

$$\Rightarrow \begin{bmatrix} 1 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

so $$x_1 + x_2 \frac{1}{\sqrt{2}} - x_4 \frac{1}{\sqrt{2}} = 0$$

$$x_2 \frac{1}{\sqrt{2}} + x_3 + x_4 \frac{1}{\sqrt{2}} = 3$$

thus $$x_1 = \frac{1}{\sqrt{2}} (x_4 - x_2)$$

$$x_3 = -\frac{1}{\sqrt{2}} (x_4 + x_2) + 3$$

Now must solve

$$\min_x \|x\|_2^2 \quad \text{s.t.} \quad \underline{\Phi} x = b$$

$$\iff \min_x \left[x_1^2 + x_2^2 + x_3^2 + x_4^2 \right] \quad (8)$$

Note $$x_1^2 = \frac{1}{2} (x_4 - x_2)^2$$

$$= \frac{1}{2} x_4^2 - x_2 x_4 + \frac{1}{2} x_2^2$$

$$x_3^2 = \left(3 - \frac{1}{\sqrt{2}} (x_2 + x_4)\right)^2$$

$$= \frac{1}{2} x_2^2 + x_2 x_4 - 3\sqrt{2}\, x_2 + \frac{1}{2} x_4^2 - 3\sqrt{2}\, x_4 + 9$$

So the optimisation in (8) becomes

$$\min_{x_2, x_4} f(x_2, x_4) \text{ where}$$

$$f(x_2, x_4) = 2x_4^2 + 2x_2^2 - 3\sqrt{2}\,x_2 - 3\sqrt{2}\,x_4 + 9$$

this is convex in both arguments so

$$\frac{\partial f}{\partial x_2}(x_2^*) = 0 = 4x_2^* - 3\sqrt{2}$$

$$\Rightarrow x_2^* = \frac{3\sqrt{2}}{4}$$

$$\frac{\partial f}{\partial x_4}(x_4^*) = 0 = 4x_4^* - 3\sqrt{2}$$

$$\Rightarrow x_4^* = \frac{3\sqrt{2}}{4}$$

Then $x_1^* = 0 \qquad x_3^* = \frac{3}{2}$

$$\therefore X^* = \begin{bmatrix} 0 \\ \frac{3\sqrt{2}}{4} \\ \frac{3}{2} \\ \frac{3\sqrt{2}}{4} \end{bmatrix}$$

Sparsity of this solution is 3.

c) Solving

$$\min_{x} \|x\|_1 \quad \text{s.t.} \quad \underline{\Phi} x = b$$

with

$$\|x\|_1 = \left|\frac{1}{\sqrt{2}}(x_4 - x_2)\right| + |x_2| + \left|3 - \frac{1}{\sqrt{2}}(x_2 + x_4)\right| + |x_4|$$

where expressions of x_1 & x_3 are as before in Part (b).

Now plot $\|x\|_1$ expression as surface or mesh in Python / Matlab gives solution at

$$x^* = \begin{bmatrix} 0 \\ 0 \\ 3 \\ 0 \end{bmatrix} \qquad \text{sparsity of } \underline{\underline{1}}$$