EE263 homework 3

1. *Linearizing range measurements.* Consider a single (scalar) measurement y of the distance or range of $x \in \mathbf{R}^n$ to a fixed point or beacon at a, *i.e.*, $y = \|x - a\|$.

   (a) Show that the linearized model near x_0 can be expressed as $\delta y = k^T \delta x$, where k is the unit vector (*i.e.*, with length one) pointing from a to x_0. Derive this analytically, and also draw a picture (for $n = 2$) to demonstrate it.

   (b) Consider the error e of the linearized approximation, *i.e.*,

   $$e = \|x_0 + \delta x - a\| - \|x_0 - a\| - k^T \delta x.$$

   The relative error of the approximation is given by $\eta = e/\|x_0 - a\|$. We know, of course, that the absolute value of the relative error is very small provided δx is small. In many specific applications, it is possible and useful to make a stronger statement, for example, to derive a bound on how large the error can be. You will do that here. In fact you will prove that

   $$0 \leq \eta \leq \frac{\alpha^2}{2}$$

   where $\alpha = \|\delta x\|/\|x_0 - a\|$ is the relative size of δx. For example, for a relative displacement of $\alpha = 1\%$, we have $\eta \leq 0.00005$, *i.e.*, the linearized model is accurate to about 0.005%. To prove this bound you can proceed as follows:

   - Show that $\eta = -1 + \sqrt{1 + \alpha^2 + 2\beta} - \beta$ where $\beta = k^T \delta x/\|x_0 - a\|$.
   - Verify that $|\beta| \leq \alpha$.
   - Consider the function $g(\beta) = -1 + \sqrt{1 + \alpha^2 + 2\beta} - \beta$ with $|\beta| \leq \alpha$. By maximizing and minimizing g over the interval $-\alpha \leq \beta \leq \alpha$ show that

   $$0 \leq \eta \leq \frac{\alpha^2}{2}.$$

Solution.

(a) For the linearized model we have

$$\delta y = \left(\frac{\partial y}{\partial x}\right) \delta x$$

so all we have to do is to compute the matrix $\partial y/\partial x$. Since $y = \|x - a\|$ we have $y^2 = (x-a)^T(x-a)$ and differentiating both sides with respect to x gives

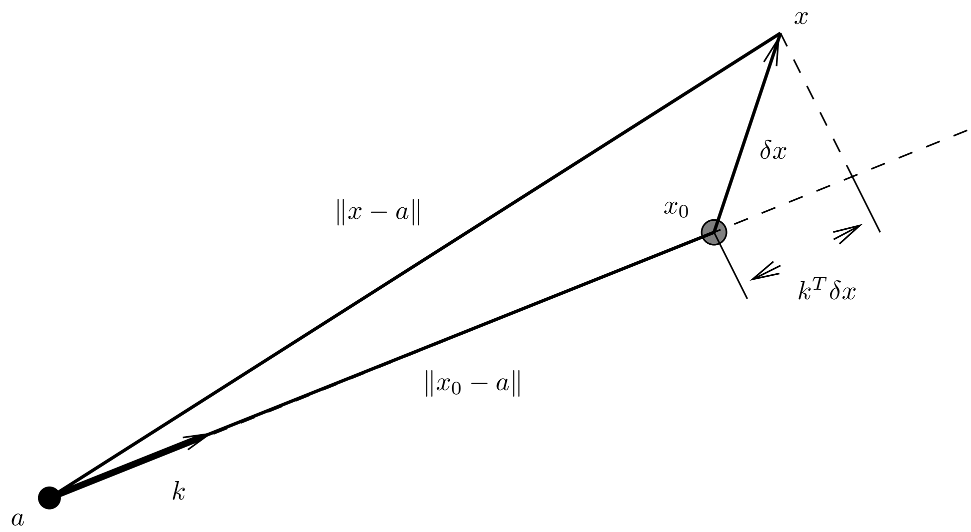
$$2\frac{\partial y}{\partial x} y = 2(x - a)^T$$

and therefore

$$\frac{\partial y}{\partial x} = \frac{(x - a)^T}{y} = \frac{(x - a)^T}{\|x - a\|},$$

so $\delta y = k^T \delta x$ with $k = (x - a)/\|x - a\|$. Clearly, k points from a to x and is of length one since

$$k^T k = \frac{(x - a)^T (x - a)}{\|x - a\|^2} = 1.$$



(b) • First we show that $\eta = -1 + \sqrt{1+\alpha^2+2\beta} - \beta$ where $\beta = k^T\delta x/\|x_0-a\|$. Note that

$$\begin{aligned} e &= \|x_0+\delta x - a\| - \|x_0-a\| - k^T\delta x \\ &= \|x_0-a\|\left(\|\frac{x_0-a}{\|x_0-a\|} + \frac{\delta x}{\|x_0-a\|}\| - 1 - \frac{k^T\delta x}{\|x_0-a\|}\right), \end{aligned}$$

and after dividing both sides by $\|x_0-a\|$ and using $k=(x_0-a)/\|x_0-a\|$, $\beta = k^T\delta x/\|x_0-a\|$ and $\eta = e/\|x_0-a\|$ we get

$$\eta = \|k + \frac{\delta x}{\|x_0-a\|}\| - 1 - \beta. \tag{1}$$

But

$$\begin{aligned} \|k+\frac{\delta x}{\|x_0-a\|}\| &= \sqrt{\left(k+\frac{\delta x}{\|x_0-a\|}\right)^T\left(k+\frac{\delta x}{\|x_0-a\|}\right)} \\ &= \sqrt{\|k\|^2 + 2\frac{k^T\delta x}{\|x_0-a\|} + \frac{\|\delta x\|^2}{\|x_0-a\|^2}}. \end{aligned}$$

Since $\|k\|=1$ and by substituting the values for α and β we have

$$\|k+\frac{\delta x}{\|x_0-a\|}\| = \sqrt{1+2\beta+\alpha^2}.$$

Therefore (**??**) can be written as

$$\eta = \sqrt{1+2\beta+\alpha^2} - 1 - \beta.$$

• It is easy to see that $|\beta| \le \alpha$. Simply we can use the Cauchy-Schwarz inequality for the vectors k and $\delta x/\|x_0-a\|$, *i.e.*,

$$\left|\frac{k^T\delta x}{\|x_0-a\|}\right| \le \|k\|\frac{\|\delta x\|}{\|x_0-a\|},$$

and since $\|k\|=1$ we immediately get $|\beta|\le\alpha$.

- At this point, all we need to do to derive a bound on how large the error can be is to maximize and minimize the function $g(\beta) = \sqrt{1+2\beta+\alpha^2} - 1 - \beta$ over the interval $|\beta| \leq \alpha$ or $-\alpha \leq \beta \leq \alpha$. The maximum or minimum of a smooth function ($g(\beta)$) over a given interval ($-\alpha \leq \beta \leq \alpha$) can only occur at the endpoints of the interval ($\beta = \pm\alpha$) or at the extremums (points β with $g'(\beta) = 0$). For $g(\beta)$ we have:
  - *Value at endpoint $\beta = \alpha$.*

$$\begin{aligned} g(\alpha) &= \sqrt{1+2\alpha+\alpha^2} - 1 - \alpha \\ &= \sqrt{(1+\alpha)^2} - 1 - \alpha \qquad \text{(note: } 1+\alpha > 0 \text{ since } \alpha \geq 0) \\ &= 1 + \alpha - 1 - \alpha \\ &= 0. \end{aligned}$$

  - *Value at endpoint $\beta = -\alpha$.*

$$\begin{aligned} g(-\alpha) &= \sqrt{1-2\alpha+\alpha^2} - 1 - \alpha \\ &= \sqrt{(1-\alpha)^2} - 1 + \alpha \\ &= |1-\alpha| - (1-\alpha) \\ &= \begin{cases} 0; & 0 \leq \alpha \leq 1 \\ 2(\alpha-1); & \alpha > 1. \end{cases} \end{aligned}$$

    Therefore $g(-\alpha) \geq 0$ for all α because $2(\alpha-1) > 0$ for $\alpha > 1$.
  - *Extremum value.*

$$g'(\beta) = \frac{1}{\sqrt{1+2\beta+\alpha^2}} - 1.$$

    Setting $g'(\beta) = 0$ we get $\sqrt{1+2\beta+\alpha^2} = 1$ or $1+2\beta+\alpha^2 = 1$ and therefore $\beta_{\text{ex.}} = -\alpha^2/2$. The function value at the extremum $\beta_{\text{ex.}} = -\alpha^2/2$ is

$$\begin{aligned} g(\beta_{\text{ex.}}) &= \sqrt{1-\alpha^2+\alpha^2} - 1 + \frac{\alpha^2}{2} \\ &= \frac{\alpha^2}{2}. \end{aligned}$$

  Clearly, $g(\beta) \geq 0$ for all β satisfying $|\beta| \leq \alpha$ because the value of $g(\beta)$ at the endpoints $\beta = \pm\alpha$ and at the extremum $\beta = \alpha^2/2$ are all non-negative. Thus we have achieved the lower bound on the relative error η, *i.e.*, we have shown that $\eta \geq 0$. For the upper bound we need to be a bit more careful. The upper bound we get is either $g(\alpha)$, $g(-\alpha)$ or $g(\beta_{\text{ex.}})$. First note that $g(\alpha) = 0$ is always less than or equal to $g(\beta_{\text{ex.}}) = \alpha^2/2 \geq 0$ so the choice of $g(\alpha)$ is immediately ruled out as the maximum of g. Now consider $g(-\alpha)$ and $g(\beta_{\text{ex.}})$. For $0 \leq \alpha \leq 1$ we obviously have $g(\beta_{\text{ex.}}) \geq g(-\alpha) = 0$. For $\alpha > 1$ we also have $g(\beta_{\text{ex.}}) \geq g(-\alpha) = 2(\alpha-1)$ because $\alpha^2/2 \geq 2(\alpha-1)$ is equivalent to $\alpha^2 - 4\alpha + 4 \geq 0$ which is true since $\alpha^2 - 4\alpha + 4 = (\alpha-2)^2$ is a complete square. Thus, we acheive an upper bound on $g(\beta)$ for all β satisfying $|\beta| \leq \alpha$ as $g(\beta) \leq \alpha^2/2$. Therefore we have shown that $\eta \leq \alpha^2/2$ and we are done.
  (Note: when $\beta_{\text{ex.}}$ falls outside the interval $\beta \leq |\alpha|$, it is possible to achieve a tighter upper bound for g. In this case, the maximum of g over $\beta \leq |\alpha|$ is obtained at the endpoint $\beta = -\alpha/2$. The extremum $\beta_{\text{ex.}} = -\alpha^2/2$ falls outside $\beta \leq |\alpha|$ when $\alpha^2/2 > \alpha$ or $\alpha > 2$. Therefore, a tighter upper bound on η for $\alpha > 2$ becomes $\eta \leq g(-\alpha) = 2(\alpha-1)$.)

2. *Halfspace.* Suppose $a, b \in \mathbf{R}^n$ are two given points. Show that the set of points in $\mathbf{R}^n$ that are closer to a than b is a halfspace, *i.e.*:

$$\{x \mid \|x-a\| \leq \|x-b\| \} = \{ x \mid c^T x \leq d\}$$

   for appropriate $c \in \mathbf{R}^n$ and $d \in \mathbf{R}$. Give c and d explicitly, and draw a picture showing a, b, c, and the halfspace.

Solution. It is easy to see geometrically what is going on: the hyperplane that goes right between a and b splits $\mathbf{R}^n$ into two parts; the points closer to a (than b) and the points closer to b (than a). More precisely, the hyperplane is normal to the line through a and b, and intersects that line at the midpoint between a and b. Now that we have the idea, let's try to derive it algebraically. Let x belong to the set of points in $\mathbf{R}^n$ that are closer to a than b. Therefore $\|x-a\| < \|x-b\|$ or $\|x-a\|^2 < \|x-b\|^2$ so

$$(x-a)^T(x-a) < (x-b)^T(x-b).$$

Expanding the inner products gives

$$x^Tx - x^Ta - a^Tx + a^Ta < x^Tx - x^Tb - b^Tx + b^Tb$$

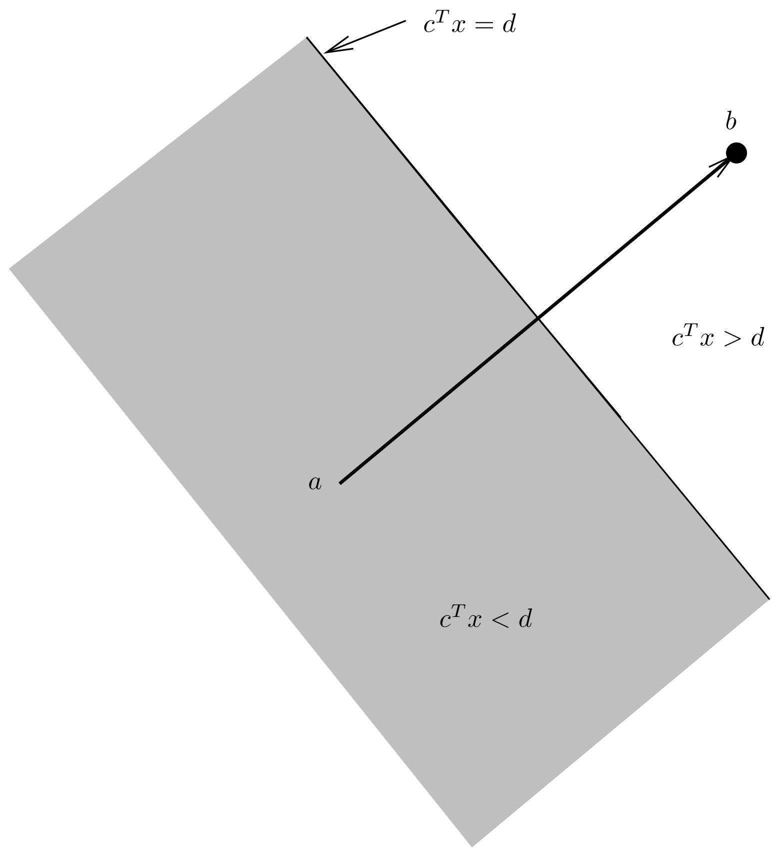
or

$$-2a^Tx + a^Ta < -2b^Tx + b^Tb$$

and finally

$$(b-a)^Tx < \frac{1}{2}(b^Tb - a^Ta). \tag{2}$$

Thus (**??**) is in the form $c^Tx < d$ with $c = b-a$ and $d = \frac{1}{2}(b^Tb - a^Ta)$ and therefore we have shown that the set of points in $\mathbf{R}^n$ that are closer to a than b is a halfspace. Note that the hyperplane $c^Tx = d$ is perpendicular to $c = b-a$.



3. *Some properties of the product of two matrices.* For each of the following statements, either show that it is true, or give a (specific) counterexample.

- If AB is full rank then A and B are full rank.
- If A and B are full rank then AB is full rank.
- If A and B have zero nullspace, then so does AB.
- If A and B are onto, then so is AB.

You can assume that $A \in \mathbf{R}^{m\times n}$ and $B \in \mathbf{R}^{n\times p}$. Some of the false statements above become true under certain assumptions on the dimensions of A and B. As a trivial example, all of the statements above are true when A and B are scalars, *i.e.*, $n = m = p = 1$. For each of the statements above, find conditions on n, m, and p that make them true. Try to find the most general conditions you can. You can give your conditions as inequalities involving n, m, and p, or you can use more informal language such as "A and B are both skinny."

Solution. First note that an $m \times n$ matrix is full rank if and only if the maximum number of independent columns or rows is equal to $\min\{m, n\}$.

- If AB is full rank then A and B are full rank. *False.* Consider the following counter example:

$$A = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad AB = \begin{bmatrix} 1 & 1 \end{bmatrix}.$$

  Clearly AB is full rank while B is not.

- If A and B are full rank then AB is full rank. *False.* Consider:

$$A = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad AB = 0.$$

  Clearly, A and B are full rank while AB is not.

- If A and B have zero null space, then so does AB. *True.* The proof is easy. We will prove that $ABx = 0$ implies that $x = 0$ and hence $\mathcal{N}(AB) = \{0\}$. If $ABx = 0$, since A has zero null space then $Bx = 0$. Now since B has zero null space this implies that $x = 0$ and we are done.
- If A and B are onto, then so is AB. *True.* We need to show that $y = ABx$ can be solved in x given any y. Suppose that $y \in \mathbf{R}^m$ is arbitrary. Since A is onto, then $y = A\tilde{x}$ holds for some $\tilde{x} \in \mathbf{R}^n$. Now consider the equation $\tilde{x} = Bx$. Since B is onto, then $\tilde{x} = Bx$ holds for some $x \in \mathbf{R}^p$. This proves that $y = ABx$ is solvable in x with $y = A\tilde{x}$ and $\tilde{x} = Bx$ and we are done.

Now we will find conditions under which the first two statements are correct. We will give these conditions based on the relative sizes of m, n and p, *i.e.*, when A is fat or skinny, B is fat or skinny, or AB is fat or skinny. We consider a square matrix to be both fat and skinny. There are 8 possible cases to check, but by using transposes we can reduce that down to 4 cases. For example lets consider the case when AB is full rank and fat, A is fat and B is fat we are considering wheiter A and B are full rank. Since AB is full rank, $(AB)^T$ will also be full rank. We know that $(AB)^T = B^T A^T$ so the same results apply for AB skinny, B and A skinny. First we consider the statement: "If AB is full rank then A and B are full rank."

- *A fat, B fat, AB fat* (or *A skinny, B skinny, AB skinny.*) The statement is not true for this case. Consider the counter example:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}}_{\text{full rank}} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}}_{\text{not full rank}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}}_{\text{full rank}}.$$

- *A fat, B skinny, AB fat* (or *A fat, B skinny, AB skinny.*) The statement is not true in this case. Consider:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}}_{\text{full rank}} \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}}_{\text{not full rank}} = \underbrace{\begin{bmatrix} 1 & 1 \end{bmatrix}}_{\text{full rank}}.$$

  However, if we add the constraint that AB is square then the statement becomes correct. To show this we use the facts that for a full rank fat matrix A all rows are independent, so $x^T A = 0$ implies $x = 0$, and for a full rank skinny matrix B all columns are independent, so $Bx = 0$ implies that $x = 0$. We first prove by contradiction that AB full rank implies that A is full rank. If A (fat) is not full rank, then there exists an $x \neq 0$ such that $x^T A = 0$, and therefore, $x^T AB = 0$. This implies that the rows of AB (a square matrix) are dependent which is impossible since AB is full rank and we are done. Now we prove that B should be full rank as well. If B (skinny) is not full rank, then $Bx = 0$ for $x \neq 0$ which implies that $ABx = 0$, or the columns of AB (a full rank square matrix) are dependent which is a contradiction. Hence B is full rank too and we are done.

- *A skinny, B fat, AB fat* (or *A skinny, B fat, AB skinny.*) The statement is true in this case. First note that if AB is full rank then A should be square. We have

$$\mathbf{rank}(AB) \leq \mathbf{min}\{\mathbf{rank}(A), \mathbf{rank}(B)\}$$

  and since A is skinny and B is fat, $\mathbf{rank}(A) \leq n$ and $\mathbf{rank}(B) \leq n$ and therefore

$$\mathbf{rank}(AB) \leq n.$$

  Now since AB is full rank and fat, then $\mathbf{rank}(AB) = m$ so $m \leq n$. However, A is skinny so $m \geq n$ and therefore we can only have $m = n$ or that A is square. Now it is easy to prove that AB full rank implies that A and B are full rank. We first prove that A is full rank by contradiction. Suppose that A (square) is not full rank so there exists a non-trivial linear combination of its rows that is equal to zero, *i.e.*, $x \neq 0$ and $x^T A = 0$. Therefore, $x^T AB = 0$ which implies that a linear combination of the rows of AB (a fat matrix) is zero which is impossible because AB is full rank. This shows that A should be full rank. Now we show that B should be full rank as well. Since A is full rank and square, then A^{-1} exists so $B = A^{-1}(AB)$. Suppose that B (fat) is not full rank so there exists an $x \neq 0$ such that $x^T B = 0$ and therefore $x^T A^{-1}(AB) = 0$. But $x^T A^{-1}$ is nonzero because x is nonzero and A^{-1} is invertible, which implies that a linear combination of the rows of AB (a full rank fat matrix) is zero. This is impossible of course and we have shown by contradiction that B should be full rank and we are done.

- *A fat, B fat, AB skinny* (or *A skinny, B skinny, AB fat.*) If A is fat, B is fat and AB is skinny, then A, B and AB can only be square matrices. A being fat implies that $m \leq n$ and B being fat implies that $n \leq p$ and we get $p \geq m$. However, $p \leq m$ because AB is skinny, so we can only have $m = p$, and therefore $m = n$ as well. In other words, A, B and AB are square. As a result, this case (A square, B square, AB square) falls into the previous category (A skinny, B fat, AB fat) and hence the statement is true.

To summarize, the most general conditions for the statement to be true are:

- A fat, B skinny, AB square,
- A square, B fat, AB fat,
- A skinny, B square, AB skinny.

Comment: Another way to do this part:
The following inequalities are always true, regardless of the sizes of A, B and AB:

$$\mathbf{rank}(A) \leq \min\{m, n\}, \quad \mathbf{rank}(B) \leq \min\{n, p\}$$

$$\mathbf{rank}(AB) \leq \min\{\mathbf{rank}(A), \mathbf{rank}(B)\}$$

Since AB is full rank, we also have $\mathbf{rank}(AB) = \mathbf{min}\{m, p\}$. From this and the last inequality above we get the following:

$$\min\{m, p\} \leq \mathbf{rank}(A) \leq \min\{m, n\}, \quad \min\{m, p\} \leq \mathbf{rank}(B) \leq \min\{n, p\}$$

Now, with the three numbers m, n and p, there are six different cases. However, as mentioned before, we only need to check three cases, since the other three can be obtained by taking transposes. Using the above inequalities in each case, we get:

- $m \leq n \leq p$: $\mathbf{rank}(A) = m, \ m \leq \mathbf{rank}(B) \leq n$
  Thus in this case A will be full rank, but we can't say anything about B. The only way to be able to infer that B is also full rank is to have $m = n$. So the claim will be true if $m = n \leq p$.
- $m \leq p \leq n$: $\mathbf{rank}(A) = m, \ m \leq \mathbf{rank}(B) \leq p$
  Similar to the previous case, to be able to infer both A and B are full rank, we should have $m = p$. So the condition in this case will be $m = p \leq n$.
- $n \leq m \leq p$: $m \leq \mathbf{rank}(A) \leq n$, but $n \leq m$, so we must have $m = n \leq p$, yielding $\mathbf{rank}(A) = \mathbf{rank}(B) = m$.

Therefore, the most general conditions where the claim is true are:
$m = n \leq p, \ n = p \leq m, \ m = p \leq n$
Which are the same conditions as the ones obtained before.
Now we consider the second statement: "If A and B are full rank then AB is full rank." Again we consider different cases:

- *A fat, B fat, AB fat* (or *A skinny, B skinny, AB skinny.*) The statement is true in this case. Since AB is fat, we need to prove that $x^T AB = 0$ implies that $x = 0$. But this is easy: $x^T AB = 0$ implies that $x^T A = 0$ (because B is fat and full rank) and $x^T A = 0$ implies that $x = 0$ (because A is fat and full rank) and we are done.
- *A fat, B skinny, AB fat* (or *A fat, B skinny, AB skinny.*) The statement is not true in this case. Consider the counter example:

$$\underbrace{\begin{bmatrix} 1 & 0 \end{bmatrix}}_{\text{full rank}} \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{\text{full rank}} = \underbrace{[0]}_{\text{not full rank}}.$$

- *A skinny, B fat, AB fat* (or *A skinny, B fat, AB skinny.*) The statement is not true in this case. Consider:

$$\underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\text{full rank}} \underbrace{\begin{bmatrix} 1 & 0 \end{bmatrix}}_{\text{full rank}} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}_{\text{not full rank}}.$$

- *A fat, B fat, AB skinny* (or *A skinny, B skinny, AB fat.*) As shown previously, if A is fat, B is fat and AB is skinny, then A, B and AB can only be square matrices. Therefore, this case falls into the category of A fat, B fat, AB fat for which the statement is true.

To summarize, the statement is true only if

- A fat, B fat, AB fat,
- A skinny, B skinny, AB skinny.

4. *Some true/false questions.* Determine if the following statements are true or false. No justification or discussion is needed for your answers. What we mean by "true" is that the statement is true for all values of the matrices and vectors given. You can't assume anything about the dimensions of the matrices (unless it's explicitly stated), but you can assume that the dimensions are such that all expressions make sense. For example, the statement "$A+B=B+A$" is true, because no matter what the dimensions of A and B (which must, however, be the same), and no matter what values A and B have, the statement holds. As another example, the statement $A^2 = A$ is false, because there are (square) matrices for which this doesn't hold. (There are also matrices for which it does hold, *e.g.*, an identity matrix. But that doesn't make the statement true.)

   a. If all coefficients (*i.e.*, entries) of the matrix A are positive, then A is full rank.

   Solution. False. The matrix $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ has all entries positive and is singular, hence not full rank.

   b. If A and B are onto, then $A+B$ must be onto.

   Solution. False. The 1×1 matrix $A = 1$ is full rank, and so is the matrix $B = -1$. But $A + B = 0$ (the 1×1 zero), which is not onto.

   c. If A and B are onto, then so is the matrix $\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$.

   Solution. True. To show this matrix is onto, we need to show that we can solve the equations

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

   for any y_1 and y_2. (These are all vectors.) The bottom block row is $y_2 = Bx_2$. Using the fact that B is onto, we can find at least one x_2 such that $y_2 = Bx_2$. The top block row is

$$y_1 = Ax_1 + Cx_2,$$

   which we can rewrite as

$$Ax_2 = Cx_2 - y_1.$$

   Using the fact that A is onto, we can find at least one x_1 that satisfies this equation. Now we're done.

   d. If A and B are onto, then so is the matrix $\begin{bmatrix} A \\ B \end{bmatrix}$.

   Solution. False. Let A and B both be the 1×1 matrix 1. These are each onto, but $\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is not.

   e. If the matrix $\begin{bmatrix} A \\ B \end{bmatrix}$ is onto, then so are the matrices A and B.

Solution. True. To say that $\begin{bmatrix} A \\ B \end{bmatrix}$ is onto means that for any vector y, we can find at least one x that satisfies

$$y = \begin{bmatrix} A \\ B \end{bmatrix} x.$$

Let's use this to show that A and B are both onto. First let's consider the equation $z = Au$. We can solve this by finding an x that satisfies

$$\begin{bmatrix} z \\ 0 \end{bmatrix} = \begin{bmatrix} A \\ B \end{bmatrix} x.$$

In a similar way can solve the equation $w = Bv$ for any vector w.

f. If A is full rank and skinny, then so is the matrix $\begin{bmatrix} A \\ B \end{bmatrix}$.

Solution. True. Since the matrix A is skinny and full rank, its has zero nullspace: whenever we have $Ax = 0$, we can conclude $x = 0$. The matrix $\begin{bmatrix} A \\ B \end{bmatrix}$ is also skinny, so to show it is full rank we must show that it, too, has zero nullspace. To do this suppose that

$$\begin{bmatrix} A \\ B \end{bmatrix} x = 0.$$

This means that $Ax = 0$ and $Bx = 0$. From the first, we conclude that $x = 0$. This shows that $\begin{bmatrix} A \\ B \end{bmatrix}$ is full rank.

5. *Orthogonal complement of a subspace.* If $\mathcal{V}$ is a subspace of $\mathbf{R}^n$ we define $\mathcal{V}^\perp$ as the set of vectors orthogonal to every element in $\mathcal{V}$, *i.e.*,

$$\mathcal{V}^\perp = \{\, x \mid \langle x, y \rangle = 0,\ \forall y \in \mathcal{V} \,\}.$$

(a) Verify that $\mathcal{V}^\perp$ is a subspace of $\mathbf{R}^n$.

(b) Suppose $\mathcal{V}$ is described as the span of some vectors $v_1, v_2, \ldots, v_r$. Express $\mathcal{V}$ and $\mathcal{V}^\perp$ in terms of the matrix $V = \begin{bmatrix} v_1 & v_2 & \cdots & v_r \end{bmatrix} \in \mathbf{R}^{n\times r}$ using common terms (range, nullspace, transpose, etc.)

(c) Show that every $x \in \mathbf{R}^n$ can be expressed uniquely as $x = v + v^\perp$ where $v \in \mathcal{V}$, $v^\perp \in \mathcal{V}^\perp$. *Hint:* let v be the projection of x on $\mathcal{V}$.

(d) Show that $\dim \mathcal{V}^\perp + \dim \mathcal{V} = n$.

(e) Show that $\mathcal{V} \subseteq \mathcal{U}$ implies $\mathcal{U}^\perp \subseteq \mathcal{V}^\perp$.

Solution.

(a) We do not need to check all properties of a vector space to hold for $\mathcal{V}^\perp$, since many of them hold only because $\mathcal{V}^\perp \subseteq \mathbf{R}^n$ and the vector sum and scalar product definitions over $\mathcal{V}^\perp$ and $\mathbf{R}^n$ are the same. We only need to verify the following properties:

- $0 \in \mathcal{V}^\perp$.
- $\forall x_1, x_2 \in \mathcal{V}^\perp :\ x_1 + x_2 \in \mathcal{V}^\perp$.
- $\forall \alpha \in \mathbf{R},\ \forall x \in \mathcal{V}^\perp :\ \alpha x \in \mathcal{V}^\perp$.

The first property comes from the fact that $\langle 0, y\rangle = 0$ for all $y \in \mathbf{R}^n$ and therefore $0 \in \mathcal{V}^\perp$. To verify the second property, we pick two arbitrary elements x_1 and x_2 in $\mathcal{V}^\perp$ and show that $x_1 + x_2 \in \mathcal{V}^\perp$. Let y be any vector in $\mathbf{R}^n$. We have

$$\begin{aligned}\langle x_1 + x_2, y\rangle &= \langle x_1, y\rangle + \langle x_2, y\rangle \\ &= 0 + 0 \qquad \text{(since } x_1 \in \mathcal{V}^\perp \text{ and } x_2 \in \mathcal{V}^\perp\text{)} \\ &= 0,\end{aligned}$$

and therefore $x_1 + x_2 \in \mathcal{V}^\perp$. Finally, we show that if $\alpha \in \mathbf{R}$ and $x \in \mathcal{V}^\perp$ then $\alpha x \in \mathcal{V}^\perp$. If $y \in \mathbf{R}^n$ is arbitrary

$$\begin{aligned}\langle \alpha x, y\rangle &= \alpha\langle x, y\rangle \\ &= \alpha \cdot 0 \qquad \text{(since } x \in \mathcal{V}^\perp\text{)} \\ &= 0,\end{aligned}$$

which by definition of $\mathcal{V}^\perp$, proves that $\alpha x \in \mathcal{V}^\perp$ and we are done.

(b) Expressing $\mathcal{V}$ in terms of the matrix V is easy. The span of vectors $v_1, v_2, \ldots, v_r$ is simply all linear combinations of the columns of V and therefore $\mathcal{V} = \mathcal{R}(V)$. To express $\mathcal{V}^\perp$ in terms of V we use the trivial fact that $x \in \mathcal{V}^\perp$ if and only if $x \perp v_i$ for $i = 1, \ldots, r$. (If $x \perp v_i$ then x is orthogonal to any linear combination of the v_i's and hence any element in $\mathcal{V}^\perp$. If $x \in \mathcal{V}^\perp$ then x is specially orthogonal to the vectors $v_i \in \mathcal{V}^\perp$ for $i = 1, \ldots, r$.) Therefore $x \in \mathcal{V}^\perp$ if and only if $v_i^T x = 0$ for $i = 1, \ldots, r$. In other words, using matrix notation, $x \in \mathcal{V}^\perp$ if and only if

$$\begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_r^T \end{bmatrix} x = 0$$

or $V^T x = 0$. Therefore $\mathcal{V}^\perp = \mathcal{N}(V^T)$.

(c) Suppose that $w_1, w_2, \ldots, w_k$ is an orthonormal basis for $\mathcal{V}$. Consider the projection of x on $\mathcal{V}$, *i.e.*,

$$v := (w_1^T x)w_1 + (w_2^T x)w_2 + \cdots + (w_k^T x)w_k.$$

Clearly, $v \in \mathcal{V}$ because it is a linear combination of the basis vectors w_i. Now we show that $x - v$ (projection error) is an element in $\mathcal{V}^\perp$. To do this we only have to verify that $x - v \perp w_i$ or $w_i^T(x - v) = 0$ for $i = 1, \ldots, k$. This is easy because

$$\begin{aligned}w_i^T(x - v) &= w_i^T x - w_i^T v \\ &= w_i^T x - (w_i^T x) w_i^T w_i \qquad \text{(since } w_i^T w_j = 0 \text{ for } i \neq j\text{)} \\ &= 0 \qquad \text{(since } w_i^T w_i = 1\text{)}.\end{aligned}$$

Now that $x - v \in \mathcal{V}^\perp$, define $v^\perp \in \mathcal{V}^\perp$ as $v^\perp = x - v$ so $x = v + v^\perp$ with $v \in \mathcal{V}$ and $v^\perp \in \mathcal{V}^\perp$. Now we show that the decomposition $x = v + v^\perp$ is unique. Suppose that there are two ways to express x as the sum of elements in $\mathcal{V}$ and $\mathcal{V}^\perp$, *i.e.*, $x = v_1 + v_1^\perp$ and $x = v_2 + v_2^\perp$ where $v_1, v_2 \in \mathcal{V}$ and $v_1^\perp, v_2^\perp \in \mathcal{V}^\perp$. Therefore $v_1 + v_1^\perp = v_2 + v_2^\perp$ or $v_1 - v_2 = v_1^\perp - v_2^\perp$. But $v_1 - v_2 \in \mathcal{V}$ (because $v_1, v_2 \in \mathcal{V}$) and $v_1^\perp - v_2^\perp \in \mathcal{V}^\perp$ (because $v_1^\perp, v_2^\perp \in \mathcal{V}^\perp$), and by definition of $\mathcal{V}^\perp$ we should have $(v_1 - v_2) \perp (v_1^\perp - v_2^\perp)$ or $(v_1 - v_2)^T(v_1^\perp - v_2^\perp) = 0$. Now since $v_1 - v_2 = v_1^\perp - v_2^\perp$ this implies that

$$(v_1 - v_2)^T(v_1 - v_2) = \|v_1 - v_2\|^2 = 0$$

and

$$(v_1^\perp - v_2^\perp)^T(v_1^\perp - v_2^\perp) = \|v_1^\perp - v_2^\perp\|^2 = 0$$

so $v_1 = v_2$ and $v_1^\perp = v_2^\perp$ or the decomposition is *unique*.

(d) This follows from the previous part. In part (**??**) we showed that any vector in $\mathbf{R}^n$ can be expressed as the sum of two elements in $\mathcal{V}$ and $\mathcal{V}^\perp$. Therefore, if $\{w_i\}_{i=1}^k$ is a basis for $\mathcal{V}$ and $\{u_i\}_{i=1}^l$ is a basis for $\mathcal{V}^\perp$, for arbitrary $x \in \mathbf{R}^n$ the scalars α_i and β_i exist such that

$$x = \sum_{i=1}^{k} \alpha_i w_i + \sum_{i=1}^{l} \beta_i u_i$$

or the set of vectors $\{w_1, \ldots, w_k, u_1, \ldots, u_l\}$ span $\mathbf{R}^n$. In fact, the vectors w_i for $i = 1, \ldots, k$ are orthogonal to the vectors u_i for $i = 1, \ldots, l$ by the definition of $\mathcal{V}^\perp$ and are therefore independent. Since the set of vectors $\{w_1, \ldots, w_k, u_1, \ldots, u_l\}$ span $\mathbf{R}^n$ and $w_1, \ldots, w_k, u_1, \ldots, u_l$ are independent we get

$$\dim \mathcal{V} + \dim \mathcal{V}^\perp = k + l = n.$$

(e) To show that $\mathcal{U}^\perp \subseteq \mathcal{V}^\perp$ we take an arbitrary element $x \in \mathcal{U}^\perp$ and prove that $x \in \mathcal{V}^\perp$. Since $x \in \mathcal{U}^\perp$ then $x \perp y$ for all $y \in \mathcal{U}$. But $\mathcal{V} \subseteq \mathcal{U}$ so we also have $x \perp y$ for all $y \in \mathcal{V}$. By definition of $\mathcal{V}^\perp$, this is nothing but to state that $x \in \mathcal{V}^\perp$ and we are done.

6. *Temperatures in a multi-core processor.* We are concerned with the temperature of a processor at two critical locations. These temperatures, denoted $T = (T_1, T_2)$ (in degrees C), are affine functions of the power dissipated by three processor cores, denoted $P = (P_1, P_2, P_3)$ (in W). We make 4 measurements. In the first, all cores are idling, and dissipate 10W. In the next three measurements, one of the processors is set to full power, 100W, and the other two are idling. In each experiment we measure and note the temperatures at the two critical locations.

P_1	P_2	P_3	T_1	T_2
10W	10W	10W	27°	29°
100W	10W	10W	45°	37°
10W	100W	10W	41°	49°
10W	10W	100W	35°	55°

Suppose we operate all cores at the same power, p. How large can we make p, without T_1 or T_2 exceeding 70°?

You must fully explain your reasoning and method, in addition to providing the numerical solution.

Solution. The temperature vector T is an affine function of the power vector P, *i.e.*, we have $T = AP + b$ for some matrix $A \in \mathbf{R}^{2\times 3}$ and some vector $b \in \mathbf{R}^2$. Once we find A and b, we can predict the temperature T for *any* value of P.

The first approach is to (somewhat laboriously) write equations describing the measurements in terms of the elements of A. Let a_{ij} denote the (i, j) entry of A. We can write out the relations $T = AP + b$ for the 4 experiments listed above as the set of 8 equations

$$\begin{aligned}
10a_{11} + 10a_{12} + 10a_{13} + b_1 &= 27,\\
10a_{21} + 10a_{22} + 10a_{23} + b_2 &= 29,\\
100a_{11} + 10a_{12} + 10a_{13} + b_1 &= 45,\\
100a_{21} + 10a_{22} + 10a_{23} + b_2 &= 37,\\
10a_{11} + 100a_{12} + 10a_{13} + b_1 &= 41,\\
10a_{21} + 100a_{22} + 10a_{23} + b_2 &= 49,\\
10a_{11} + 10a_{12} + 100a_{13} + b_1 &= 35,\\
10a_{21} + 10a_{22} + 100a_{23} + b_2 &= 55.
\end{aligned}$$

Next, we define a vector of unknowns, $x = (a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, b_1, b_2) \in \mathbf{R}^8$. We rewrite the 8 equations above as $Cx = d$, where

$$C = \begin{bmatrix} 10 & 10 & 10 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 10 & 10 & 10 & 0 & 1 \\ 100 & 10 & 10 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 100 & 10 & 10 & 0 & 1 \\ 10 & 100 & 10 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 10 & 100 & 10 & 0 & 1 \\ 10 & 10 & 100 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 10 & 10 & 100 & 0 & 1 \end{bmatrix}, \qquad d = \begin{bmatrix} 27 \\ 29 \\ 45 \\ 37 \\ 41 \\ 49 \\ 35 \\ 55 \end{bmatrix}.$$

We solve for x as $x = C^{-1}d$. (It turns out that C is invertible.) Putting the entries of x into the appropriate places in A and b, we have

$$A = \begin{bmatrix} 0.200 & 0.156 & 0.089 \\ 0.089 & 0.222 & 0.289 \end{bmatrix}, \qquad b = \begin{bmatrix} 22.6 \\ 23.0 \end{bmatrix}.$$

At this point we can predict T for any P (assuming we trust the affine model).

Substituting $P = (p, p, p)$ into $T = AP + b$, we get

$$T_1 = 0.444p + 22.6, \qquad T_2 = 0.600p + 23.0.$$

Both of these temperatures are increasing in p (it would be quite surprising if this were not the case). The value of p for which $T_1 = 70$ is $p = (70 - 22.6)/0.444 = 106.8$W. The value of p for which $T_2 = 70$ is $p = (70 - 23)/0.6 = 78.3$W. Thus, the maximum value of p for which both temperatures do not exceed $70°$ is $p = 78.3$W.

Alternative solution. Another way of solving this problem is to directly exploit the fact that T is an affine function of P. This means that if we form any linear combination of the power vectors used in the experiment, *with the coefficients summing to one*, the temperature vector will also be the same linear combination of the temperatures.

By averaging the last three experiments we find if the powers are $P = (40, 40, 40)$, then the temperature vector is $T = (40.33, 47.00)$. (Note that this is really a prediction, based on the observed experimental data and the affineness assumption; it's not a new experiment!)

Now we form a new power vector of the form

$$P = (1 - \theta)(10, 10, 10) + \theta(40, 40, 40) = (10 + 30\theta, 10 + 30\theta, 10 + 30\theta),$$

where $\theta \in \mathbf{R}$. The coefficients $1-\theta$ and θ sum to one, so since T is affine, we find that the corresponding temperature vector is

$$T = (1 - \theta)(27, 29) + \theta(40.33, 47.00) = (27 + 13.33\theta, 29 + 18\theta),$$

just as above. The first coefficient hits 70 at $\theta = 3.22$; the second coefficient hits 70 at $\theta = 2.23$. Thus, θ can be as large as $\theta = 2.27$. This corresponds to the powers $P = (78.3, 78.3, 78.3)$.