Homework 2 Solutions

EE 263 Stanford University Summer 2018

July 5, 2018

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

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

for appropriate $c \in \mathbb{R}^n$ and $d \in \mathbb{R}$. 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 $\mathbb{R}^n$ 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 $\mathbb{R}^n$ that are closer to a than b. Therefore $\|x-a\| < \|x-b\|$ or $\|x-a\|^2 < \|x-b\|^2$ so

$$(x-a)^{\mathsf{T}}(x-a) < (x-b)^{\mathsf{T}}(x-b).$$

Expanding the inner products gives

$$x^{\mathsf{T}}x - x^{\mathsf{T}}a - a^{\mathsf{T}}x + a^{\mathsf{T}}a < x^{\mathsf{T}}x - x^{\mathsf{T}}b - b^{\mathsf{T}}x + b^{\mathsf{T}}b$$

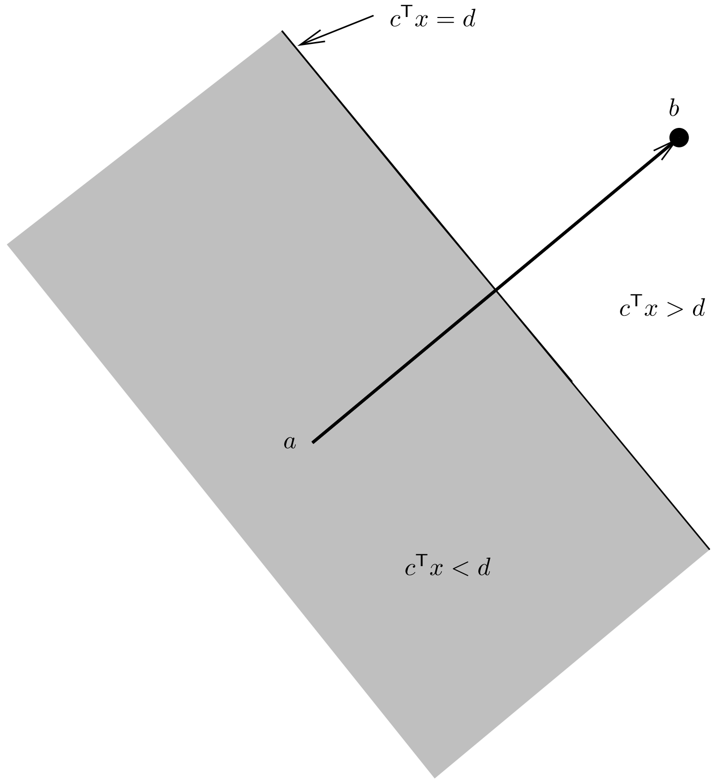
or

$$-2a^{\mathsf{T}}x + a^{\mathsf{T}}a < -2b^{\mathsf{T}}x + b^{\mathsf{T}}b$$

and finally

$$(b-a)^{\mathsf{T}}x < \frac{1}{2}(b^{\mathsf{T}}b - a^{\mathsf{T}}a). \tag{1}$$

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



2. **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 $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$. 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 $\text{null}(AB) = \{0\}$. 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 $y \in \mathbb{R}^m$ is arbitrary. Since A is onto, then $y = A\tilde{x}$ holds for some $\tilde{x} \in \mathbb{R}^n$. Now consider the equation $\tilde{x} = Bx$. Since B is onto, then $\tilde{x} = Bx$ holds for some $x \in \mathbb{R}^p$. 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, $(AB)^\mathsf{T}$ will also be full rank. We know that $(AB)^\mathsf{T} = B^\mathsf{T} A^\mathsf{T}$ 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 $x^\mathsf{T} A = 0$ 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

$x^\mathsf{T} A = 0$, and therefore, $x^\mathsf{T} AB = 0$. 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 $x^\mathsf{T} A = 0$. Therefore, $x^\mathsf{T} AB = 0$ 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 $x^\mathsf{T} B = 0$ and therefore $x^\mathsf{T} A^{-1}(AB) = 0$. But $x^\mathsf{T} A^{-1}$ 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 $x^\mathsf{T} AB = 0$ implies that $x = 0$. But this is easy: $x^\mathsf{T} AB = 0$ implies that $x^\mathsf{T} A = 0$ (because B is fat and full rank) and $x^\mathsf{T} A = 0$ 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.

3. **Geometric analysis of navigation.** Let $(x, y) \in \mathbb{R}^2$ be the unknown coordinates of a point in the plane, let $(p_i, q_i) \in \mathbb{R}^2$ be the known coordinates of a beacon for $i = 1, \ldots, n$, and let ρ_i be the measured distance between (x, y) and the ith beacon. The linearized navigation equations near a point $(x_0, y_0) \in \mathbb{R}^2$ are

$$\delta\rho \approx A \begin{bmatrix} \delta x \\ \delta y \end{bmatrix},$$

where we define the matrix $A \in \mathbb{R}^{n \times 2}$ such that

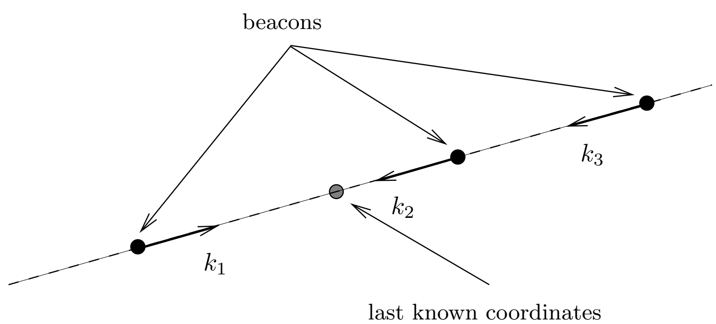
$$A_{i1} = \frac{x_0 - p_i}{\|(x_0, y_0) - (p_i, q_i)\|} \qquad \text{and} \qquad A_{i2} = \frac{y_0 - q_i}{\|(x_0, y_0) - (p_i, q_i)\|}.$$

Find the conditions under which A has full rank. Describe the conditions geometrically (*i.e.*, in terms of the relative positions of the unknown coordinates and the beacons).

Solution. Suppose we have n beacons in $\mathbb{R}^2$ and the linearized equation is $y = Ax$ so $y \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times 2}$ and $x \in \mathbb{R}^2$, where

$$A = \begin{bmatrix} k_1^\mathsf{T} \\ k_2^\mathsf{T} \\ \vdots \\ k_n^\mathsf{T} \end{bmatrix}$$

and where k_i is the unit vector pointing from the ith beacon to the last known position. Since $A \in \mathbb{R}^{n \times 2}$, A is full rank if and only if at least two rows of A, or equivalently, two of the unit vectors k_i are independent. Independence of two vectors means that one should not be a scalar multiple of the other. Therefore, the only case for which A is *not* full rank is when all vectors k_i are parallel. Geometrically, in terms of the relative positions of the last known coordinates and the beacons, this means that the beacons and the last known coordinates all lie on the same line (see figure). If the beacons and the last known coordinates do not lie on the same line then A is full rank.

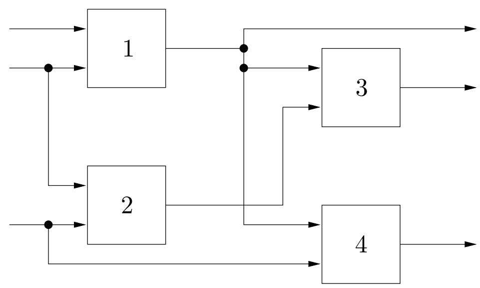


4. **Digital circuit gate sizing.** A digital circuit consists of a set of n (logic) gates, interconnected by wires. Each gate has one or more inputs (typically between one and four), and one output, which is connected via the wires to other gate inputs and possibly to some external circuitry. When the output of gate i is connected to an input of gate j, we say that gate i *drives* gate j, or that gate j is in the *fan-out* of gate i. We describe the topology of the circuit by the *fan-out list* for each gate, which tells us which other gates the output of a gate connects

to. We denote the fan-out list of gate i as $\mathrm{FO}(i) \subseteq \{1,\dots,n\}$. We can have $\mathrm{FO}(i) = \emptyset$, which means that the output of gate i does not connect to the inputs of any of the gates $1,\dots,n$ (presumably the output of gate i connects to some external circuitry). It's common to order the gates in such a way that each gate only drives gates with higher indices, *i.e.*, we have $\mathrm{FO}(i) \subseteq \{i+1,\dots,n\}$. We'll assume that's the case here. (This means that the gate interconnections form a directed acyclic graph.)

To illustrate the notation, a simple digital circuit with $n = 4$ gates, each with 2 inputs, is shown below. For this circuit we have

$$\mathrm{FO}(1) = \{3,4\}, \quad \mathrm{FO}(2) = \{3\}, \quad \mathrm{FO}(3) = \emptyset, \quad \mathrm{FO}(4) = \emptyset.$$



The 3 input signals arriving from the left are called *primary inputs*, and the 3 output signals emerging from the right are called *primary outputs* of the circuit. (You don't need to know this, however, to solve this problem.)

Each gate has a (real) *scale factor* or *size* x_i. These scale factors are the design variables in the gate sizing problem. They must satisfy $1 \le x_i \le x^{\max}$, where $x^{\max}$ is a given maximum allowed gate scale factor (typically on the order of 100). The total area of the circuit has the form

$$A = \sum_{i=1}^{n} a_i x_i,$$

where a_i are positive constants.

Each gate has an *input capacitance* C_i^{in}, which depends on the scale factor x_i as

$$C_i^{\mathrm{in}} = \alpha_i x_i,$$

where α_i are positive constants.

Each gate has a *delay* d_i, which is given by

$$d_i = \beta_i + \gamma_i C_i^{\mathrm{load}} / x_i,$$

where β_i and γ_i are positive constants, and C_i^{load} is the *load capacitance* of gate i. Note that the gate delay d_i is always larger than β_i, which can be intepreted as the minimum possible delay of gate i, achieved only in the limit as the gate scale factor becomes large.

The load capacitance of gate i is given by

$$C_i^{\mathrm{load}} = C_i^{\mathrm{ext}} + \sum_{j \in \mathrm{FO}(i)} C_j^{\mathrm{in}},$$

where $C_i^{\rm ext}$ is a positive constant that accounts for the capacitance of the interconnect wires and external circuitry.

We will follow a simple design method, which assigns an equal delay T to all gates in the circuit, *i.e.*, we have $d_i = T$, where $T > 0$ is given. For a given value of T, there may or may not exist a feasible design (*i.e.*, a choice of the x_i, with $1 \le x_i \le x^{\rm max}$) that yields $d_i = T$ for $i = 1, \ldots, n$. We can assume, of course, that $T > \max_i \beta_i$, *i.e.*, T is larger than the largest minimum delay of the gates.

Finally, we get to the problem.

a) Explain how to find a design $x^\star \in \mathbb{R}^n$ that minimizes T, subject to a given area constraint $A \le A^{\rm max}$. You can assume the fanout lists, and all constants in the problem description are known; your job is to find the scale factors x_i. Be sure to explain how you determine if the design problem is feasible, *i.e.*, whether or not there is an x that gives $d_i = T$, with $1 \le x_i \le x^{\rm max}$, and $A \le A^{\rm max}$.

   Your method can involve any of the methods or concepts we have seen so far in the course. It can also involve a simple search procedure, *e.g.*, trying (many) different values of T over a range.

   Note: this problem concerns the general case, and not the simple example shown above.

b) Carry out your method on the particular circuit with data given in the file `gate_sizing_data.json`

   The fan-out lists are given as an $n \times n$ matrix `F`, with i, j entry one if $j \in \mathrm{FO}(i)$, and zero otherwise. In other words, the ith row of `F` gives the fanout of gate i. The jth entry in the ith row is 1 if gate j is in the fan-out of gate i, and 0 otherwise.

Comment. You do not need to know anything about digital circuits; *everything* you need to know is stated above.

Solution.

a) We define the fanout matrix F as $F_{ij} = 1$, if $j \in \mathrm{FO}(i)$, and $F_{ij} = 0$ otherwise. The matrix F is *strictly upper triangular*, since $\mathrm{FO}(i) \subseteq \{i+1, \ldots, n\}$.

   Using the formulas given above, and $d_i = T$, we have

$$\begin{aligned} T &= d_i \\ &= \beta_i + \gamma_i \frac{C_i^{\rm load}}{x_i} \\ &= \beta_i + \gamma_i \frac{C_i^{\rm ext} + \sum_{j \in \mathrm{FO}(i)} C_j^{\rm in}}{x_i} \\ &= \beta_i + \gamma_i \frac{C_i^{\rm ext} + \sum_{j \in \mathrm{FO}(i)} \alpha_j x_j}{x_i}. \end{aligned}$$

   Multiplying by x_i we get the equivalent equations

$$T x_i = \beta_i x_i + \gamma_i \left(C_i^{\rm ext} + \sum_{j \in \mathrm{FO}(i)} \alpha_j x_j \right),$$

which we can express in matrix form as

$$Tx = \mathrm{diag}(\beta)x + \mathrm{diag}(\gamma)C^{\mathrm{ext}} + \mathrm{diag}(\gamma)F\,\mathrm{diag}(\alpha)x.$$

Defining

$$K = \mathrm{diag}(\beta) + \mathrm{diag}(\gamma)F\,\mathrm{diag}(\alpha),$$

we can write the equations as

$$(TI - K)x = \mathrm{diag}(\gamma)C^{\mathrm{ext}},$$

a set of n linear equations in n unknowns. So this problem really does belong in EE263, after all.

For choices of T for which $TI - K$ is nonsingular, there is only one solution of this set of linear equations,

$$x = (TI - K)^{-1}\,\mathrm{diag}(\gamma)C^{\mathrm{ext}}.$$

If this x happens to satisfy $1 \le x_i \le x^{\max}$, and $A = a^{\mathsf{T}}x \le A^{\max}$, then it is a feasible design. Our job, then, is to find the smallest T for which this occurs. If it occurs for no T, then the problem is infeasible.

Let's analyze the issue of singularity of $TI - K$. The matrix K is upper triangular, with diagonal elements β_i. So $TI - K$ is upper triangular, with diagonal elements $T - \beta_i$. But these are all positive, by our assumption. So the matrix $TI - K$ is nonsingular.

Thus, for each value of T (larger than $\max_i \beta_i$) there is exactly one possible choice of gate sizes. Among the ones that are feasible, we have to choose the one corresponding to the smallest value of T.

We can solve this problem by examing a reasonable range of values of T, and for each value, finding x. We check whether x is feasible, by looking at $\min_i x_i$, $\max_i x_i$, and A. We take our final design as the one which is feasible, and has smallest value of T. Alternatively, we can start with a value of T just a little bit larger than $\max_i \beta_i$, then increase T until we find our first feasible x, which we take as our solution.

b) The following code generates x for a range of value of T, and plots $\min_i x_i$, $\max_i x_i$, and A, versus T.

```
gate_sizing_data

deltaT=0.001;
Trange=max(beta)+deltaT:deltaT:6;
i=1;
for T=Trange
    K=diag(beta)+diag(gamma)*F*diag(alpha);
    x=(T*eye(n)-K)\diag(gamma)*Cext;
    maxX(i)=max(x);
    minX(i)=min(x);
    Area(i)=a'*x;
    i=i+1;
```

```
end

res=Area<=Amax & minX>=1 & (maxX<=xmax);
index=find(res);
T=Trange(index(1))

subplot(3,1,1)
plot(Trange,minX)
ylabel('minx')
axis([2 6 0 4])
line([2,6],[1,1],'Color','r')
grid on
subplot(3,1,2)
plot(Trange,maxX)
ylabel('maxx')
axis([2 6 0 150])
line([2,6],[100,100],'Color','r')
grid on
subplot(3,1,3)
plot(Trange,Area)
xlabel('T')
ylabel('A')
axis([2 6 0 500])
line([2,6],[400,400],'Color','r')
grid on
```

The output of the code is

```
T= 2.5194
```

Figure **??** shows how the minimum and maximum gate sizes, and the total area, vary with T, with the blue lines showing the limits. This shows that the feasible designs correspond to $2.5194 \leq T \leq 5.088$.

A few more comments about this problem:

- Since the matrix $TI - K$ is upper triangular, we can solve for x very, very quickly. In fact, if we use sparse matrix operations, we can easily compute x very quickly (seconds or less) for a problem with $n = 10^5$ gates or more. You didn't need to know this; we're just pointing it out for fun.
- The plots above show that as T increases, all of gate sizes decrease. This implies that $\min_i x_i$, $\max_i x_i$, and A all decrease as T increases. This means you can use a more efficient bisection search to find the optimal T. Again, you didn't need to know this; we're just pointing it out.

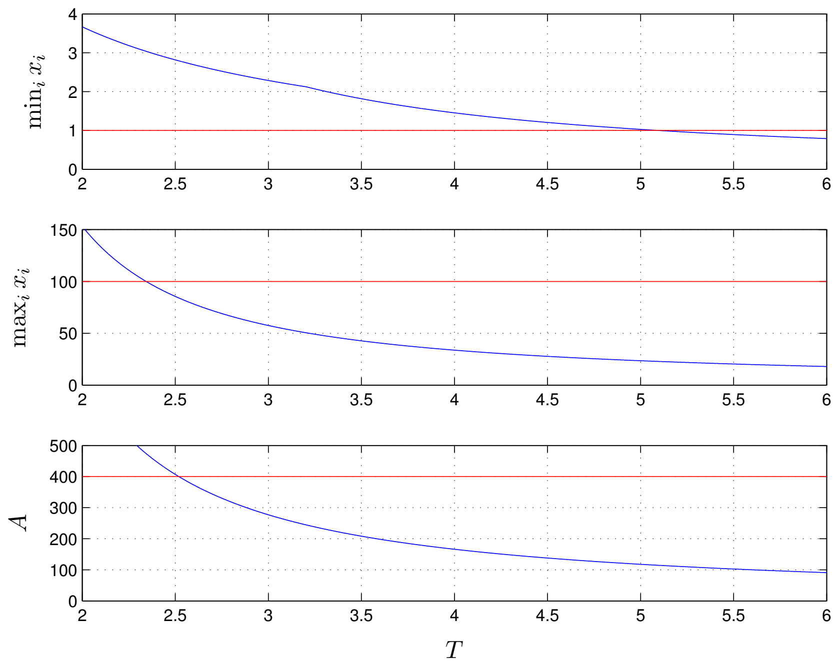


Figure 1: $\max_i x_i, \min_i x_i$, and A versus T.

5. Projection matrices. A matrix $P \in \mathbb{R}^{n \times n}$ is called a *projection matrix* if $P = P^\mathsf{T}$ and $P^2 = P$.

a) Show that if P is a projection matrix then so is $I - P$.

b) Suppose that the columns of $U \in \mathbb{R}^{n \times k}$ are orthonormal. Show that UU^T is a projection matrix. (Later we will show that the converse is true: every projection matrix can be expressed as UU^T for some U with orthonormal columns.)

c) Suppose $A \in \mathbb{R}^{n \times k}$ is full rank, with $k \leq n$. Show that $A(A^\mathsf{T} A)^{-1} A^\mathsf{T}$ is a projection matrix.

d) If $S \subseteq \mathbb{R}^n$ and $x \in \mathbb{R}^n$, the point y in S closest to x is called the *projection of x on S*. Show that if P is a projection matrix, then $y = Px$ is the projection of x on range(P). (Which is why such matrices are called projection matrices ...)

Solution.

a) To show that $I - P$ is a projection matrix we need to check two properties:

   i. $I - P = (I - P)^\mathsf{T}$

   ii. $(I - P)^2 = I - P$.

   The first one is easy: $(I - P)^\mathsf{T} = I - P^\mathsf{T} = I - P$ because $P = P^\mathsf{T}$ (P is a projection matrix.) The show the second property we have

$$
\begin{aligned}
(I-P)^2 &= I - 2P + P^2 \\
&= I - 2P + P \qquad\qquad (\text{since } P = P^2) \\
&= I - P
\end{aligned}
$$

and we are done.

b) Since the columns of U are orthonormal we have $U^\top U = I$. Using this fact it is easy to prove that $UU^\top$ is a projection matrix, *i.e.*, $(UU^\top)^\top = UU^\top$ and $(UU^\top)^2 = UU^\top$. Clearly, $(UU^\top)^\top = (U^\top)^\top U^\top = UU^\top$ and

$$
\begin{aligned}
(UU^\top)^2 &= (UU^\top)(UU^\top) \\
&= U(U^\top U)U^\top \\
&= UU^\top \qquad\qquad (\text{since } U^\top U = I).
\end{aligned}
$$

c) First note that $\left(A(A^\top A)^{-1}A^\top\right)^\top = A(A^\top A)^{-1}A^\top$ because

$$
\begin{aligned}
\left(A(A^\top A)^{-1}A^\top\right)^\top &= (A^\top)^\top \left((A^\top A)^{-1}\right)^\top A^\top \\
&= A\left((A^\top A)^\top\right)^{-1} A^\top \\
&= A(A^\top A)^{-1}A^\top.
\end{aligned}
$$

Also $\left(A(A^\top A)^{-1}A^\top\right)^2 = A(A^\top A)^{-1}A^\top$ because

$$
\begin{aligned}
\left(A(A^\top A)^{-1}A^\top\right)^2 &= \left(A(A^\top A)^{-1}A^\top\right)\left(A(A^\top A)^{-1}A^\top\right) \\
&= A\left((A^\top A)^{-1}A^\top A\right)(A^\top A)^{-1}A^\top \\
&= A(A^\top A)^{-1}A^\top \qquad\qquad (\text{since } (A^\top A)^{-1}A^\top A = I).
\end{aligned}
$$

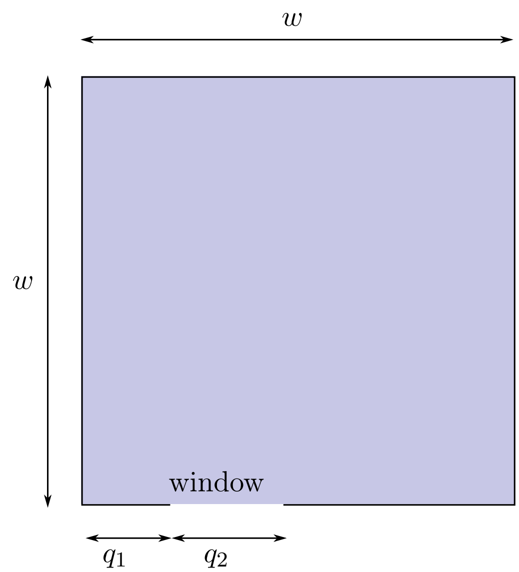
d) To show that Px is the projection of x on range(P) we verify that the "error" $x - Px$ is orthogonal to *any* vector in range(P). Since range(P) is nothing but the span of the columns of P we only need to show that $x - Px$ is orthogonal to the columns of P, or in other words, $P^\top(x - Px) = 0$. But

$$
\begin{aligned}
P^\top(x - Px) &= P(x - Px) \qquad\qquad (\text{since } P = P^\top) \\
&= Px - P^2x \\
&= 0 \qquad\qquad (\text{since } P^2 = P)
\end{aligned}
$$

and we are done.

6. **Light reflection from diffuse surfaces.** We consider a simple model for the reflection of light from a surface. The idea of the model is that the amount of energy flux leaving a surface is the sum of the amount produced by the surface (if it is a light emitter) plus the amount reflected from the surface. To simplify matters, we'll assume that the surfaces are *Lambertian*, that is, a single ray of light incident on the surface is reflected equally in all directions, so no matter where we view the surface from it will look equally bright. This is a good model for surfaces that are not shiny.

You are an architect, charged with determining how the light in a room with one window will look like. We'll work in a two-dimensional world (which simplifies much of the drudgery of computing meshes). A horizontal slice through the room is below.



The room has four walls, and the south wall has a window in it, of width q_2, placed at a distance q_1 from the west wall. We divide each wall up into segments called elements. The total light flux from element i, called the *radiosity* of i, will be denoted by x_i. This light is emitted in all directions, and the light flux from element j that hits element i is given by $F_{ij}x_j$. The number F_{ij} is called the *form factor*, and only depends on the geometry. It turns out that there is a simple formula for the form factor, given by the *strings rule*. (You don't need to be able to prove this.) Specifically,

$$F_{ij} = \frac{AD + BC - AC - BD}{2l_i}$$

where l_i is the length of element i, and AD denotes the length of the line from point A to point D, as shown in the figure below. That is, F_{ij} is the sum of the lengths of the 'crossed strings' minus the sum of the lengths of the 'uncrossed strings', divided by $2l_i$. Since all light must go somewhere, we have

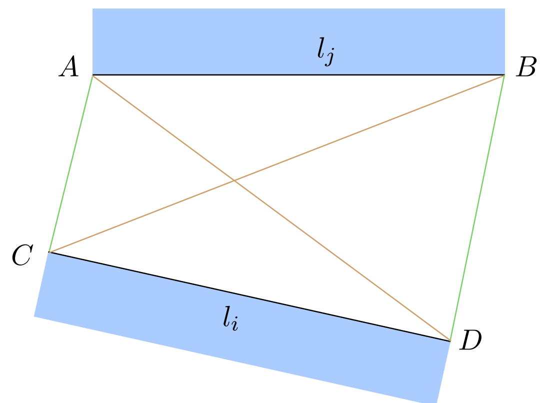
$$\sum_i F_{ij} = 1$$

where F_{ii} is defined to be zero. The light reflected from element i is then

$$\rho_i \sum_j F_{ij} x_j$$

where ρ_i is the reflectivity coefficient of the element i. The window does not reflect light, but it emits light, with a flux of 1 unit per element. Light enters through the window, bounces around the room, and the walls are therefore illuminated nonuniformly depending on position.

We will have $w = 100$, $q_1 = 20$, $q_2 = 20$, and $\rho_i = 0.8$ for the room walls. We will use elements of length 1, so in total we will divide the walls (plus window) of the room up into 400 elements.



a) Let x be the vector of radiosities of each of the elements. What is the equation that determines the radiosity x?

b) Find x, and plot the radiosity of each of the walls as a function of position along the wall.

c) Which point on the walls is the brightest (excluding the window)? What is the radiosity there?

Solution.

a) Define $\rho_i = 0$ for elements i which are part of the window. Let $R = \mathrm{diag}(\rho)$. Let the emitted light vector y be

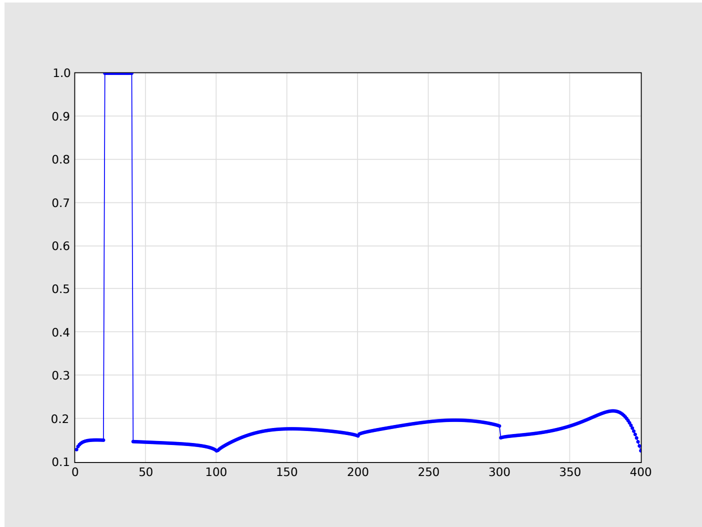
$$y_i = \begin{cases} 1 & \text{for } i \text{ in the window} \\ 0 & \text{otherwise} \end{cases}$$

Then we have

$$x = y + RFx$$

It turns out that in this case, this equation has a unique solution.

b) The plot is below.



Here the elements are number anticlockwise going around the room, starting at the southwest corner.

c) The brightest point is 20m north along the west wall from the southwest corner. It has radiosity 0.218.

7. True/false questions about linear algebra. Determine whether each of the following statements is true or false. In each case, give either a proof or a counterexample.

a) If Q has orthonormal columns, then $\|Q^\mathsf{T} w\| \le \|w\|$ for all vectors w.

b) Suppose $A \in \mathbb{R}^{m\times p}$ and $B \in \mathbb{R}^{m\times q}$. If $\operatorname{null}(A) = \{0\}$ and $\operatorname{range}(A) \subset \operatorname{range}(B)$, then $p \le q$.

c) If $V = \begin{bmatrix} V_1 & V_2 \end{bmatrix}$ is invertible and $\operatorname{range}(V_1) = \operatorname{null}(A)$, then $\operatorname{null}(AV_2) = \{0\}$.

d) If $\operatorname{rank}\left(\begin{bmatrix} A & B \end{bmatrix}\right) = \operatorname{rank}(A) = \operatorname{rank}(B)$, then $\operatorname{range}(A) = \operatorname{range}(B)$.

e) Suppose $A \in \mathbb{R}^{m\times n}$. Then, $x \in \operatorname{null}(A^\mathsf{T})$ if and only if $x \notin \operatorname{range}(A)$.

f) Suppose A is invertible. Then, AB is not full rank if and only if B is not full rank.

g) If A is not full rank, then there is a nonzero vector x such that $Ax = 0$.

Solution.

a) The statement is true. Suppose $Q \in \mathbb{R}^{m\times n}$. Because the columns of Q are orthonormal, and hence linearly independent, we know that $m \ge n$. If $m = n$, then Q is an orthogonal matrix, so QQ^T, and we have that

$$\|Q^\mathsf{T} w\|^2 = w^\mathsf{T}(QQ^\mathsf{T})w = w^\mathsf{T} w = \|w\|^2.$$

Now consider the case when $m > n$. Let $q_1, \ldots, q_n$ be the columns of Q:

$$Q = \begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix}.$$

Then, we can extend $(q_1, \ldots, q_n)$ to an orthonormal basis $(q_1, \ldots, q_m)$ for $\mathbb{R}^m$. Define the matrix $\tilde{Q} \in \mathbb{R}^{m \times (m-n)}$ such that

$$\tilde{Q} = \begin{bmatrix} q_{m+1} & \cdots & q_n \end{bmatrix}.$$

Then, we have that $\begin{bmatrix} Q & \tilde{Q} \end{bmatrix}$ is an orthogonal matrix, so that

$$\|w\|^2 = \left\| \begin{bmatrix} Q & \tilde{Q} \end{bmatrix}^\mathsf{T} w \right\|^2 = \left\| \begin{bmatrix} Q^\mathsf{T} \\ \tilde{Q}^\mathsf{T} \end{bmatrix} w \right\|^2 = \|Q^\mathsf{T} w\|^2 + \|\tilde{Q}^\mathsf{T} w\|^2 \geq \|Q^\mathsf{T} w\|^2.$$

Combining these results, we see that if the columns of Q are orthonormal, then $\|Q^\mathsf{T} w\| \leq \|w\|$ for all vectors w. (This result is known as Bessel's inequality.)

b) The statement is true. Because $\operatorname{range}(A) \subset \operatorname{range}(B)$, we have that

$$\operatorname{rank}(A) = \dim(\operatorname{range}(A)) \leq \dim(\operatorname{range}(B)) = \operatorname{rank}(B).$$

Since the rank of a matrix is bounded by its number of columns, we have that

$$\operatorname{rank}(B) \leq q.$$

Conservation of dimension tells us that

$$\operatorname{rank}(A) = \dim(\mathbb{R}^p) - \dim(\operatorname{null}(A)) = \dim(\mathbb{R}^p) - \dim(\{0\}) = p - 0 = p.$$

Combining these results, we have that

$$p = \operatorname{rank}(A) \leq \operatorname{rank}(B) \leq q.$$

c) The statement is true. Suppose $y \in \operatorname{null}(AV_2)$. Then, $V_2 y \in \operatorname{null}(A) = \operatorname{range}(V_1)$, so there exists a vector x such that $V_1 x = V_2 y$. Therefore, we have that

$$V_1 x - V_2 y = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \begin{bmatrix} x \\ -y \end{bmatrix} = 0.$$

Because $\begin{bmatrix} V_1 & V_2 \end{bmatrix}$ is invertible, this implies that

$$\begin{bmatrix} x \\ -y \end{bmatrix} = 0,$$

and hence that $y = 0$. Thus, we see that $\operatorname{null}(AV_2) = \{0\}$.

d) The statement is true. It is sufficient to show that if $\operatorname{rank}(\begin{bmatrix} A & B \end{bmatrix}) = \operatorname{rank}(A)$, then $\operatorname{range}(A) = \operatorname{range}(\begin{bmatrix} A & B \end{bmatrix})$. This implies that if $\operatorname{rank}(\begin{bmatrix} A & B \end{bmatrix}) = \operatorname{rank}(A) = \operatorname{rank}(B)$, then

$$\operatorname{range}(A) = \operatorname{range}(\begin{bmatrix} A & B \end{bmatrix}) = \operatorname{range}(B).$$

Consider any $y \in \operatorname{range}(A)$. There exists a vector x such that $Ax = y$. Then, we have that

$$\begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = Ax = y,$$

so that $y \in \operatorname{range}\big(\begin{bmatrix} A & B \end{bmatrix}\big)$. Thus, we have that $\operatorname{range}(A) \subset \operatorname{range}\big(\begin{bmatrix} A & B \end{bmatrix}\big)$. (Note that this result holds for any matrices A and B.) Now suppose that $\operatorname{rank}(A) = \operatorname{rank}\big(\begin{bmatrix} A & B \end{bmatrix}\big)$. Let $(q_1, \ldots, q_m)$ be a basis for $\operatorname{range}(A)$. Since $\operatorname{range}(A)$ is a subspace of $\operatorname{range}\big(\begin{bmatrix} A & B \end{bmatrix}\big)$, we can extend this basis to a basis $(q_1, \ldots, q_n)$ for $\operatorname{range}\big(\begin{bmatrix} A & B \end{bmatrix}\big)$. However, it must be the case that

$$n = \operatorname{rank}\big(\begin{bmatrix} A & B \end{bmatrix}\big) = \operatorname{rank}(A) = m.$$

Thus, $(q_1, \ldots, q_m)$ is a basis for both $\operatorname{range}(A)$ and $\operatorname{range}(\begin{bmatrix} A & B \end{bmatrix})$. Therefore, for any $y \in \operatorname{range}(\begin{bmatrix} A & B \end{bmatrix})$, there exist scalars $c_1, \ldots, c_m$ such that

$$y = c_1 q_1 + \cdots + c_m q_m.$$

Because $(q_1, \ldots, q_m)$ is also a basis for $\operatorname{range}(A)$, this implies that $y \in \operatorname{range}(A)$. This shows that if $\operatorname{rank}(A) = \operatorname{rank}\big(\begin{bmatrix} A & B \end{bmatrix}\big)$, then $\operatorname{range}\big(\begin{bmatrix} A & B \end{bmatrix}\big) \subset \operatorname{range}(A)$, and thereby completes the proof.

e) The statement is false. In fact, neither direction of the equivalence is true. Consider the matrix

$$A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

For $x = 0$, we have that $x \in \operatorname{null}(A^\mathsf{T})$, and $x \in \operatorname{range}(A)$; for $x = (1, 1)$, we have that $x \notin \operatorname{range}(A)$, and $x \notin \operatorname{null}(A^\mathsf{T})$.

f) The statement is true. We claim that if A is invertible, then $\operatorname{rank}(AB) = \operatorname{rank}(B)$. This implies that AB is full rank if and only if B is full rank. Let $(q_1, \ldots, q_k)$ be a basis for $\operatorname{range}(B)$. We claim that $(Aq_1, \ldots, Aq_k)$ is a basis for AB. Consider any $y \in \operatorname{range}(AB)$. There exists a vector x such that $(AB)x = y$. Note that Bx is a vector in $\operatorname{range}(B)$, so there exist scalars $c_1, \ldots, c_k$ such that

$$Bx = c_1 q_1 + \cdots + c_k q_k.$$

Then, we have that

$$y = (AB)x = A(Bx) = A(c_1 q_1 + \cdots + c_k q_k) = c_1(Aq_1) + \cdots + c_k(Aq_k).$$

This shows that $(Aq_1, \ldots, Aq_k)$ spans $\operatorname{range}(AB)$. (This is true for any matrices A and B.) Suppose there exist scalars $d_1, \ldots, d_k$ such that

$$d_1(Aq_1) + \cdots + d_k(Aq_k) = A(d_1 q_1 + \cdots + d_k q_k) = 0.$$

Since A is invertible, this implies that

$$d_1 q_1 + \cdots + d_k q_k = A^{-1} 0 = 0.$$

Then, because $(q_1, \ldots, q_k)$ is a basis, and hence linearly independent, we have that

$$d_1 = \cdots = d_k = 0.$$

This shows that $(Aq_1, \ldots, Aq_k)$ is linearly independent, and completes the proof that $(Aq_1, \ldots, Aq_k)$ is a basis for range(AB). Since the dimension of a subspace is the number of vectors in a basis for the subspace, we have that

$$\text{rank}(B) = \dim(\text{range}(B)) = k = \dim(\text{range}(AB)) = \text{rank}(AB).$$

g) The statement is true. If A is strictly fat, then there exists a nonzero vector x such that $Ax = 0$ irrespective of whether of A is full rank. Suppose A is skinny (or square). If A is not full rank, then the columns of A must be linearly dependent: that is, there exist scalars $x_1, \ldots, x_n$, at least one of which is nonzero, such that

$$x_1 A_{*1} + \cdots + x_n A_{*n} = 0.$$

In matrix form, this equation says that

$$\begin{bmatrix} A_{*1} & \cdots & A_{*n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = Ax = 0,$$

where $x \neq 0$.