Homework 4 Solutions

Due Thursday 10/22.

1. ***A Pythagorean inequality for the matrix norm.***

   Suppose that $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{p\times n}$. Show that

   $$\left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\| \le \sqrt{\|A\|^2 + \|B\|^2}.$$

   Under what conditions do we have equality?

   Solution.

   Suppose that $v \in \mathbb{R}^n$ with $\|v\| = 1$ is the maximum gain direction of $\begin{bmatrix} A \\ B \end{bmatrix}$. Hence

   $$\left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\| = \left\| \begin{bmatrix} A \\ B \end{bmatrix} v \right\|.$$

   But

   $$\left\| \begin{bmatrix} A \\ B \end{bmatrix} v \right\|^2 = \left\| \begin{bmatrix} Av \\ Bv \end{bmatrix} \right\|^2 = \|Av\|^2 + \|Bv\|^2 \le \|A\|^2 + \|B\|^2 \qquad (1)$$

   since in general the maximum gain directions of A and B are different from the maximum direction of $\begin{bmatrix} A \\ B \end{bmatrix}$ (equality holds when maximum gain input directions of A and B are the same.) Therefore we have shown that

   $$\left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\| \le \sqrt{\|A\|^2 + \|B\|^2}.$$

   Equality holds when the maximum gain direction of A and B are the same. If $v_1 \in \mathbb{R}^n$ with $\|v_1\| = 1$ is the maximum gain direction of both A and B then

   $$\left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|^2 \ge \left\| \begin{bmatrix} A \\ B \end{bmatrix} v_1 \right\|^2 = \left\| \begin{bmatrix} Av_1 \\ Bv_1 \end{bmatrix} \right\|^2 = \|Av_1\|^2 + \|Bv_1\|^2 = \|A\|^2 + \|B\|^2,$$

   and since also $\left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|^2 \le \|A\|^2 + \|B\|^2$ we should have $\left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|^2 = \|A\|^2 + \|B\|^2$. Conversely, if

   $$\left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\| = \sqrt{\|A\|^2 + \|B\|^2}$$

   then according to (**??**) the maximum gain directions of A and B should be equal to the maximum gain direction of $\begin{bmatrix} A \\ B \end{bmatrix}$, and therefore, the maximum gain directions of A and B should be equal themselves.

2. ***Eigenvalues and singular values of a symmetric matrix.***

   Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues, and let $\sigma_1, \dots, \sigma_n$ be the singular values of a matrix $A \in \mathbb{R}^{n\times n}$, which satisfies $A = A^T$. (The singular values are based on the full SVD: If $\mathrm{rank}(A) < n$, then some of the singular values are zero.) You can assume the eigenvalues (and of course singular values) are sorted, *i.e.*, $\lambda_1 \ge \cdots \ge \lambda_n$ and $\sigma_1 \ge \cdots \ge \sigma_n$. How are the eigenvalues and singular values related?

   Solution.

   Since A is symmetric it can be diagonalized using an orthogonal matrix Q as

   $$A = Q \left[\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \right] Q^T$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A. Suppose that the columns of Q are ordered such that $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$. Thus

$$A = Q \left[\begin{bmatrix} \mathrm{sgn}\lambda_1 & & \\ & \ddots & \\ & & \mathrm{sgn}\lambda_n \end{bmatrix} \right] \left[\begin{bmatrix} |\lambda_1| & & \\ & \ddots & \\ & & |\lambda_n| \end{bmatrix} \right] Q^T$$

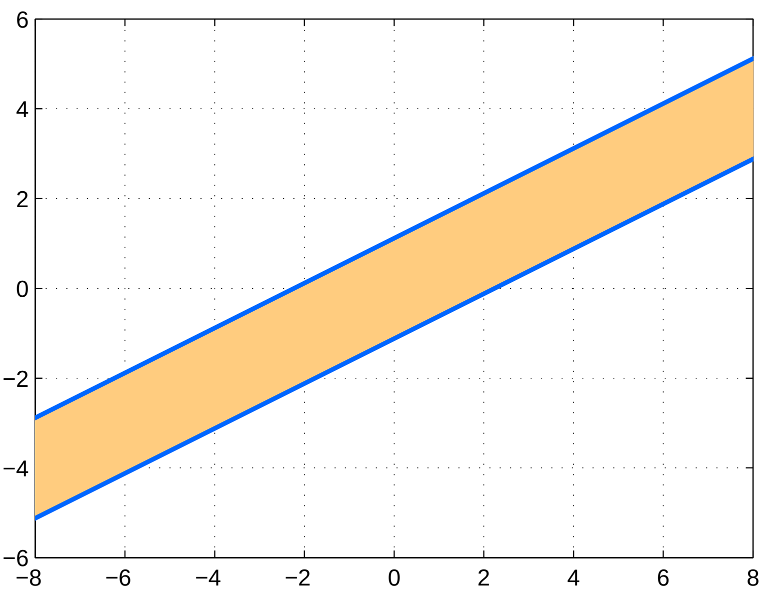
Now we define

$$U = Q \left[\begin{bmatrix} \mathrm{sgn}\lambda_1 & & \\ & \ddots & \\ & & \mathrm{sgn}\lambda_n \end{bmatrix} \right], \quad \Sigma = \left[\begin{bmatrix} |\lambda_1| & & \\ & \ddots & \\ & & |\lambda_n| \end{bmatrix} \right], \quad V = Q.$$

Clearly, U is a an orthogonal matrix because $UU^T = QQ^T = I$. Now $A = U\Sigma V^T$is a SVD of A, and we conclude that $\sigma_i = |\lambda_i|$.

3. ***Degenerate ellipsoids***

The picture below shows a degenerate ellipsoid.



In two dimensions, a degenerate ellipsoid is a *slab*; the sides are parallel lines. For the above example the slab has half-width 1 (i.e., it has width 2) and the center axis points in the direction

$$v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

We'll call the slab S.

(a) Find a symmetric matrix $Q \in \mathbb{R}^{2\times 2}$ such that the slab above is

$$S = \left\{ x \in \mathbb{R}^2 \mid x^T Q x \leq 1 \right\}$$

(b) Is Q positive definite?

(c) Consider the matrix

$$P = \begin{bmatrix} 0.04 & 0.06 \\ 0.06 & 0.09 \end{bmatrix}$$

Plot the slab corresponding to P. (Sketch it by hand, if you prefer.) What is the axis of the slab? (i.e., the long axis). What is the half-width of the slab?

(d) Now let's consider the three dimensional case. Suppose T is a cylinder, of radius 5, whose axis of symmetry passes through the origin and points along the vector

$$q = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

Find a symmetric matrix A such that the cylinder is

$$T = \left\{ x \in \mathbb{R}^3 \mid x^T A x \leq 1 \right\}$$

(e) Just for fun (i.e., not worth any points.) Is A unique? If so prove it; if not, find two different matrices A which both generate the cylinder T above.

Solution.

(a) The slab is just the set of $x \in \mathbb{R}^2$ such that

$$(w^T x)^2 \leq 1$$

where w is a unit vector orthogonal to v, i.e.,

$$w = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

Then we have

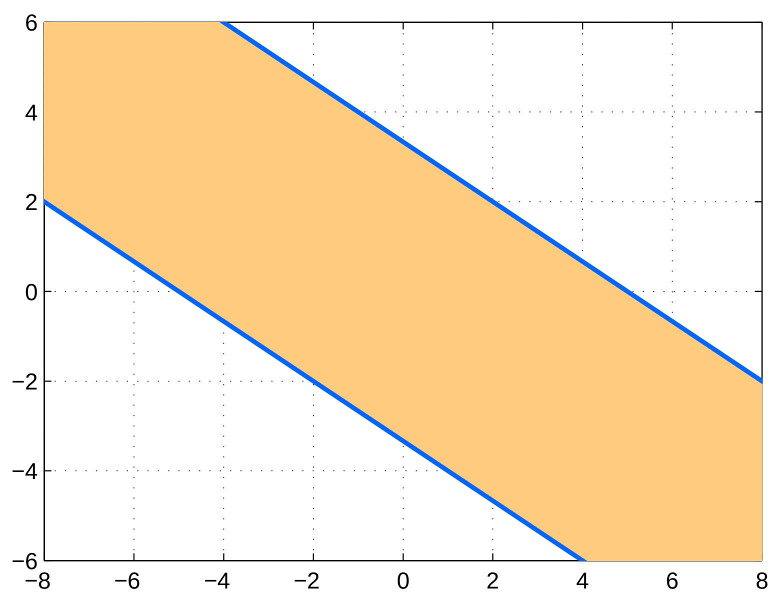
$$(w^T x)^2 = x^T w w^T x$$

and so we let $Q = ww^T$, which gives

$$Q = \begin{bmatrix} 0.2 & -0.4 \\ -0.4 & 0.8 \end{bmatrix}$$

(b) No. From above, the eigenvalues of Q are 1 and 0, so Q is only positive semidefinite.

(c) The slab is below.



The eigenvector corresponding to the non-zero eigenvalue of P is

$$v_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

with eigenvalue $\lambda = 10/\sqrt{13}$. Hence the slab half-width is $r = 1/\sqrt{\lambda} \approx 2.77$ and the slab axis is

$$v_2 = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$$

(d) The cylinder is the set of x such that

$$\|Bx\| \leq 5$$

where B is the matrix that projects onto the plane perpendicular to q. We can write this as

$$x^T A x \leq 1$$

where $A = B^T B/25$. Since the projection of a vector x onto the line in the direction r is

$$rr^T x$$

when r is a unit vector, the projection matrix is

$$B = I - \frac{qq^T}{q^T q} = \frac{1}{9}\begin{bmatrix} 8 & -2 & -2 \\ -2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}$$

and hence

$$A = \frac{1}{225}\begin{bmatrix} 8 & -2 & -2 \\ -2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}$$

Notice that $B^T B = B$ since B is an orthogonal projector.

(e) Yes, A is unique. A proof is as follows. Suppose $A_1 \neq A_2$, but the sets

$$T_1 = \left\{ x \in \mathbb{R}^3 \mid x^T A_1 x \leq 1 \right\} \qquad T_2 = \left\{ x \in \mathbb{R}^3 \mid x^T A_1 x \leq 1 \right\}$$

are equal. We must have $A_1 - A_2$ has at least one non-zero eigenvalue, and we can assume that it is positive (otherwise swap A_1 and A_2). Let v be the corresponding eigenvector, and choose $c > 0$ such that $x = cv$ satisfies

$$x^T A_2 x \leq 1$$

but

$$x^T A_1 x > 1$$

We can do this because

$$x^T A_1 x - x^T A_2 x = x^T (A_1 - A_2) x > 0$$

Then $x \in T_1$ but $x \notin T_2$, contradicting the original assumption that the sets were equal.

4. ***Sensitivity of the force to changes in the position.***

Suppose masses m_1, m_2, m_3, m_4 are located at positions x_1, x_2, x_3, x_4 in a line and connected by springs with spring constants k_{12}, k_{23}, k_{34} whose natural lengths of extension are l_{12}, l_{23}, l_{34}. Let f_1, f_2, f_3, f_4 denote the rightward forces on the m asses, e.g., $f_1 = k_{12}(x_2 - x_1 - l_{12})$.

Write the 4×4 matrix relating the column vectors f and x. Let K denote the matrix in this equation, so that

$$f = Kx + c$$

and we can measure the forces f.

Suppose the vector of positions x changes from x to $x + \delta x$, with a corresponding change in the measured force from f to $f + \delta f$. Let $v_1, \dots, v_n$ be the right singular vectors of K. Show that if $\delta x \in \operatorname{span}\{v_{k+1}, \dots, v_n\}$ then $\|\delta f\| \le \sigma_{k+1} \|\delta x\|$.

Suppose the spring constants $k_{ij} = 1$. What displacement of the position vector δx causes the least change to the forces? Can you interpret this?

Solution.

We have $f + \delta f = K(x + \delta x) + c$, so

$$\delta f = K\delta x$$

Since $\delta x \in \operatorname{span}\{v_{k+1}, \dots, v_n\}$, we must have

$$\delta x = c_{k+1}v_{k+1} + \cdots + c_n v_n$$

for some coefficients $c_{k+1}, \dots, c_n$. Therefore

$$\begin{aligned} K\delta x &= U\Sigma V^T x \\ &= \sigma_{k+1}c_{k+1}u_{k+1} + \cdots + \sigma_n c_n u_n \end{aligned}$$

where $\sigma_{k+1} \ge \cdots \ge \sigma_n \ge 0$. Therefore

$$\begin{aligned} \|\delta f\|^2 &= (K\delta x)^T K\delta x \\ &= \sum_{i=k+1}^{n} \sigma_i^2 c_i^2 \\ &\le \sigma_{k+1}^2 \sum_{i=k+1}^{n} c_i^2 \\ &= \sigma_{k+1}^2 \|\delta x\|^2 \end{aligned}$$

In this case, the matrix K is

$$K = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

Since this matrix is symmetric, its singular values are just the absolute values of its eigenvalues. So its singular values are

$$3.4142,\ 2,\ 0.5858,\ 0$$

The displacement which causes the least change to the forces is the right singular vector of K corresponding to the smallest singular value, which is

$$\delta x = \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{bmatrix}$$

Physically, this is just an equal displacement applied to all of the masses. With this displacement their relative positions do not change, so the springs exert no additional forces.

5. ***Computing the SVD by hand.***

Consider the matrix

$$A = \begin{bmatrix} -2 & 11 \\ -10 & 5 \end{bmatrix}$$

(a) Determine, on paper, an SVD of A in the usual form $A = U\Sigma V^T$. The SVD is not unique, so find the one that has the minimal number of minus signs in U and V.

(b) List the singular values, left singular vectors, and right singular vectors of A. Draw a careful, labeled picture of the unit ball in $\mathbb{R}^2$ and its image under A, together with the singular vectors, with the coordinates of their vertices marked.

(c) What is the matrix norm $\|A\|$ and the Frobenius norm $\|A\|_F$?

(d) Find A^{-1}, not directly, but via the SVD.

(e) Find the eigenvalues, λ_1 and λ_2 of A.

(f) Verify that $\det A = \lambda_1\lambda_2$ and $|\det A| = \sigma_1\sigma_2$

(g) What is the area of the ellipsoid onto which A maps the unit ball of $\mathbb{R}^2$?

Solution.

(a) The singular values of A are the non-zero eigenvalues of AA^T, which is

$$A = \begin{bmatrix} -2 & 11 \\ -10 & 5 \end{bmatrix}$$

Solving for the eigenvalues as usual gives

$$\lambda_1 = 200 \qquad \lambda_2 = 50$$

The corresponding eigenvectors

$$u_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \qquad u_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix}$$

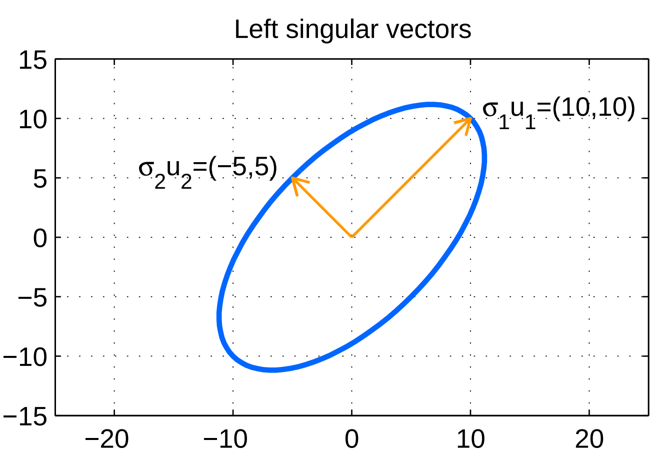
are the left singular vectors of A.
Similarly the right singular vectors are the eigenvectors of A^TA, which are

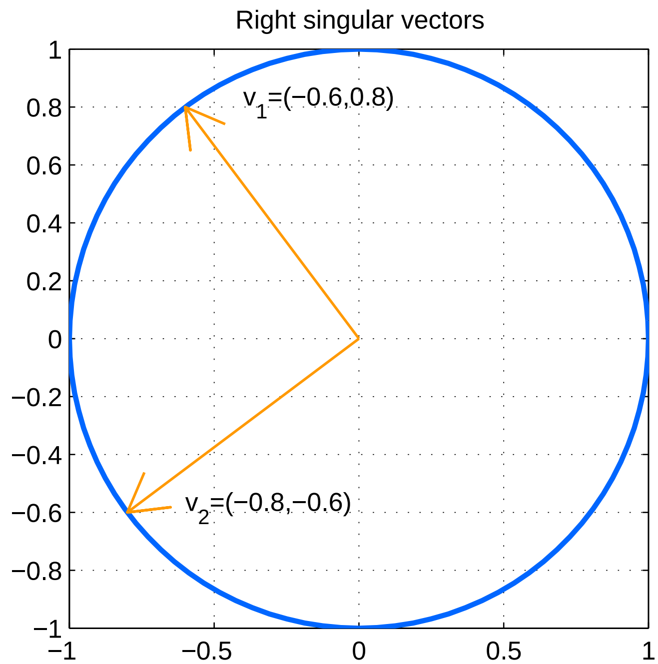
$$v_1 = \begin{bmatrix} \frac{-3}{5} \\ \frac{4}{5} \end{bmatrix} \qquad v_2 = \begin{bmatrix} \frac{4}{5} \\ \frac{3}{5} \end{bmatrix}$$

So the SVD of A is

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 10\sqrt{2} & 0 \\ 0 & 5\sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{-3}{5} & \frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix}.$$

(b) The plots are shown below.





(c) We have

$$\|A\| = 10\sqrt{2}$$
$$\|A\|_F = 5\sqrt{10}$$

(d) The inverse of A is

$$A^{-1} = V\Sigma^{-1}U^T$$

which is

$$A^{-1} = \begin{bmatrix} \frac{-3}{5} & \frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix} \begin{bmatrix} \frac{1}{10\sqrt{2}} & 0 \\ 0 & \frac{1}{5\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix} = \frac{1}{100} \begin{bmatrix} 5 & -11 \\ 10 & -2 \end{bmatrix}$$

(e) The eigenvalues of A are

$$\lambda_1 = \frac{3}{2} + \frac{1}{2}\sqrt{391}i \qquad \lambda_2 = \frac{3}{2} - \frac{1}{2}\sqrt{391}i$$

(f) We find

$$\det(A) = 100$$
$$\lambda_1\lambda_2 = \frac{9}{4} + \frac{391}{4} = 100$$
$$\sigma_1\sigma_2 = 10\sqrt{2} \times 5\sqrt{2} = 100$$

(g) We have, for any set $S \subset \mathcal{R}^n$

$$\det(A) = \frac{\text{volume}(AS)}{\text{volume}(S)}$$

So if S is the unit ball, then AS is the ellipsoid, and

$$\text{area of ellipsoid} = \det(A) \times \text{area of unit ball} = 100 \times \pi r^2 = 100\pi$$

6. ***Using the SVD for numerical computations.***

Use the SVD and Matlab to solve the following problems. Turn in your code.

(a) Find bases for the range space and null space of each of the following matrices:

$$A_1 = \begin{bmatrix} 4 & 1 & -1 \\ 3 & 2 & 0 \\ 1 & 1 & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(b) For the linear equation

$$\begin{bmatrix} 2 & -1 \\ -3 & 3 \\ -1 & 2 \end{bmatrix} x = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

is there a solution? If so find it, and say if it is unique.

(c) Find all solutions of

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

Solution.

(a) Looking at the singular values of A_1, we see that A_1 has full rank, so $\mathrm{range}(A_1) = \mathbb{R}^3$, and the standard basis will do. Since it is square, it must be invertible and we must have $\mathrm{null}(A_1) = \{0\}$.

The matrix A_2 is also full rank, and since it is fat we have $\mathrm{range}(A_2) = \mathbb{R}^3$, and again the standard basis will do. The null space is given by the span of the fourth right singular vector, i.e.,

$$\mathrm{null}(A_2) = \mathrm{span}\{v_4\}$$

We have

$$v_4 = \begin{bmatrix} -0.4082 \\ 0.8165 \\ -0.4082 \\ 0 \end{bmatrix}$$

We can also write this as

$$\mathrm{null}(A_2) = \mathrm{span}\left\{ \begin{bmatrix} -1 \\ 2 \\ -1 \\ 0 \end{bmatrix} \right\}$$

(b) Let the SVD of A be $A = U\Sigma V^T$. Then the component of $y \in \mathrm{range}(A)$ is UU^Ty, and the residual is

$$r = (I - UU^T)y$$

Computing the residual, we find that it is zero, so there is a solution to this equation. Also A is skinny and full rank (since both its singular values are nonzero), so it has zero null space and therefore the solution is unique. It is given by

$$x = \begin{bmatrix} 2 & -1 \\ -3 & 3 \\ -1 & 2 \end{bmatrix}^\dagger \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

(c) We would like to find all solutions to the equation $Ax = y$, where A and y are

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad y = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

Again, using the SVD we see that A is fat and full rank, so its range is the whole space and there is at least one solution. We can find a solution just by computing the minimum norm one, which is

$$x_{\mathrm{mn}} = A^{\dagger} y \approx \begin{bmatrix} -0.8333 \\ -0.3333 \\ 0.1667 \\ 1 \end{bmatrix}$$

The set of all solutions is then

$$\left\{ x_{\mathrm{mn}} + z \mid z \in \mathrm{null}(A) \right\}$$

In this case A has rank 3, so $\mathrm{null}(A) = \mathrm{span}\{v_4\}$, where v_4 is as in part (a). So all solutions have the form

$$x = x_{\mathrm{mn}} + cv_4$$

for some c.

7. ***Determining the number of signal sources.***

The signal transmitted by n sources is measured at m receivers. The signal transmitted by each of the sources at sampling period k, for $k = 1, \ldots, p$, is denoted by an n-vector $x(k) \in \mathbb{R}^n$. The gain from the j-th source to the i-th receiver is denoted by $a_{ij} \in \mathbb{R}$. The signal measured at the receivers is then

$$y(k) = A\,x(k) + v(k), \quad k = 1, \ldots, p,$$

where $v(k) \in \mathbb{R}^m$ is a vector of sensor noises, and $A \in \mathbb{R}^{m \times n}$ is the matrix of source to receiver gains. However, we do not know the gains a_{ij}, nor the transmitted signal $x(k)$, nor even the number of sources present n. We only have the following additional *a priori* information:

- We expect the number of sources to be less than the number of receivers (*i.e.*, $n < m$, so that A is skinny);
- A is full-rank and well-conditioned;
- All sources have roughly the same average power, the signal $x(k)$ is unpredictable, and the source signals are unrelated to each other; Hence, given enough samples (*i.e.*, p large) the vectors $x(k)$ will 'point in all directions';
- The sensor noise $v(k)$ is small relative to the received signal $A\,x(k)$.

Here's the question:

(a) You are given a large number of vectors of sensor measurements $y(k) \in \mathbb{R}^m$, $k = 1, \ldots, p$. How would you estimate the number of sources, n? Be sure to clearly describe your proposed method for determining n, and to explain when and why it works.

(b) Try your method on the signals given in the file `nsources.m`. Running this script will define the variables:

- `m`, the number of receivers;
- `p`, the number of signal samples;

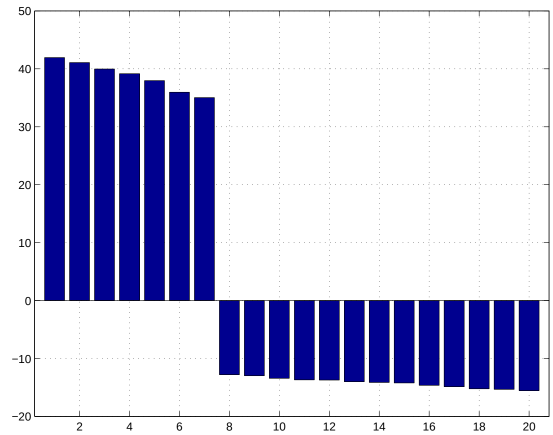


Figure 1: Singular values of Y (in dB)

- `Y`, the receiver sensor measurements, an array of size `m` by `p` (the k-th column of `Y` is $y(k)$.)

What can you say about the number of signal sources present?

Note: Our problem description and assumptions are not precise. An important part of this problem is to explain your method, and clarify the assumptions.

Solution.

The quick answer to this problem is: count the number of large singular values of Y. Since the singular values of Y are the square-root of the eigenvalues of YY^T, and YY^T is a much smaller matrix, the most efficient way to solve this problem is to find the eigenvalues of YY^T (the SVD of YY^T works as well.) With the Matlab commands:

```
z=eig(Y*Y');
stem(sqrt(z))
```

we see there are 7 large eigenvalues, which correspond to signal sources, and 13 small ones, due to noise (see figure). We conclude there are 7 sources present. Let's justify our method. There are many ways to think about this problem, we'll discuss one of them. First, collect the source signal vectors $x(k)$ in the matrix $X \in \mathbb{R}^{n\times p}$,

$$X = [\, x(1)\ x(2)\ \cdots\ x(p)\,]$$

and the noise vectors in the matrix $V \in \mathbb{R}^{m\times p}$,

$$V = [\, v(1)\ v(2)\ \cdots\ v(p)\,].$$

Now the equations can be summarized as $Y = AX + V$. We start by noting that the range of A is in an n-dimensional subspace of $\mathbb{R}^m$. Also, our assumptions mean that the matrix X is full-rank and well-conditioned: Since the $x(k)$ point equally likely in all directions, we have that

$$\frac{1}{p}\sum_{k=1}^{p} x(k)^T v$$

has about the same value for any unit length vector v (and for large enough p.) Hence, the matrix X^T has about the same gain in all directions, which implies that all the singular values of X have about the same value (and $\kappa(X)$ is on the order of 1.) Since X can map a vector into any point in $\mathbb{R}^n$, the matrix AX has the same rank and range as A, *i.e.*, can map into any point in the range of A. Also, the ratio between the maximum singular value of AX and the smallest non-zero singular value is not large (it's bounded by $\kappa(A)\,\kappa(X)$.) A geometrical interpretation is that both A and AX map a ball into an ellipsoid that's flat in $m-n$ directions, and has a moderate aspect ratio in the other directions (because of well-conditioning in the case of A, and because of the aspect ratio of the non-zero axes in the case of AX.) Another way to think about it is to view the columns of X as "filling up" a ball in $\mathbb{R}^n$. Then, the columns of AX "fill up" a flat ellipsoid in $\mathbb{R}^m$ (with n non-zero axes.) The columns of $Y = AX + V$ will no longer be in the range of A because of the noise (in fact, Y can be expected to be full-rank.) The ellipsoid is no longer flat, but not by much – if the noise is small, the semi-axes that do not correspond to the range of A are small. Another way to see this is as follows. Since AX has n large singular values, the gain of Y in the direction of the input singular vectors of AX is also large (greater or equal than $\|AXv_i\| - \|Vv_i\|$.) Also, there are $p-n$ orthogonal directions for which AX has zero gain (the nullspace of AX), and the gain of Y in those directions is small (equal to the gain of V.) In summary, Y must have n orthogonal directions with large gain, and $p-n$ orthogonal directions with small gain.