Midterm exam solutions

1. Point of closest convergence of a set of lines. We have \( m \) lines in \( \mathbb{R}^n \), described as
\[
\mathcal{L}_i = \{ p_i + tv_i \mid t \in \mathbb{R} \}, \quad i = 1, \ldots, m,
\]
where \( p_i \in \mathbb{R}^n \), and \( v_i \in \mathbb{R}^n \), with \( \|v_i\| = 1 \), for \( i = 1, \ldots, m \). We define the distance of a point \( z \in \mathbb{R}^n \) to a line \( \mathcal{L} \) as
\[
\text{dist}(z, \mathcal{L}) = \min \{ \| z - u \| \mid u \in \mathcal{L} \}.
\]
(In other words, \( \text{dist}(z, \mathcal{L}) \) gives the closest distance between the point \( z \) and the line \( \mathcal{L} \).)

We seek a point \( z^* \in \mathbb{R}^n \) that minimizes the sum of the squares of the distances to the lines,
\[
\sum_{i=1}^{m} \text{dist}(z, \mathcal{L}_i)^2.
\]
The point \( z^* \) that minimizes this quantity is called the point of closest convergence.

(a) Explain how to find the point of closest convergence, given the lines (i.e., given \( p_1, \ldots, p_m \) and \( v_1, \ldots, v_m \)). If your method works provided some condition holds (such as some matrix being full rank), say so. If you can relate this condition to a simple one involving the lines, please do so.

(b) Find the point \( z^* \) of closest convergence for the lines with data given in the Matlab file line_conv_data.m. This file contains \( n \times m \) matrices \( P \) and \( V \) whose columns are the vectors \( p_1, \ldots, p_m \), and \( v_1, \ldots, v_m \), respectively. The file also contains commands to plot the lines and the point of closest convergence (once you have found it). Please include this plot with your solution.

Solution.

(a) There are several ways to solve this problem. Our first solution starts by working out an explicit expression for \( \text{dist}(z, \mathcal{L}_i) \). To find this distance we need to solve the simple least-squares problem of minimizing \( \| z - p_i - tv_i \|^2 \) over \( t \in \mathbb{R} \). The optimal \( t \) is given by \( t^* = v_i^T(z - p_i) \), so we have
\[
\text{dist}(z, \mathcal{L}_i) = \| z - p_i - t^*v_i \| = \| (I - v_i v_i^T)(z - p_i) \|.
\]
This makes sense: we recognize \( I - v_i v_i^T \) as projection onto the orthogonal complement of the line through the origin in the direction \( v_i \), \( i.e., \) projection onto the plane with normal vector \( v_i \).

We can now set up our problem as a standard least-squares problem. We define

\[
A = \begin{bmatrix}
I - v_1 v_1^T \\
\vdots \\
I - v_m v_m^T
\end{bmatrix}, \quad b = \begin{bmatrix}
(I - v_1 v_1^T) p_1 \\
\vdots \\
(I - v_m v_m^T) p_m
\end{bmatrix},
\]

so we can write

\[
\sum_{i=1}^m \text{dist}(z, \mathcal{L}_i)^2 = \| A z - b \|^2.
\]

Now we can solve the problem, assuming \( A \) is full rank (we’ll come back to this). The solution is

\[
z^* = (A^T A)^{-1} A^T b = \left( m I - \sum_{i=1}^m v_i v_i^T \right)^{-1} \sum_{i=1}^m (p_i - v_i v_i^T p_i).
\]

Finally, let’s look at the conditions under which \( A \) is not full rank. Each \( n \times n \) block of \( A \), \( i.e., \) \( I - v_i v_i^T \), has rank exactly \( n - 1 \), with nullspace span\((v_i)\). So unless all the \( v_i \) are aligned (\( i.e., v_i = v_j \) or \( v_i = -v_j \) for all \( i, j \)), \( A \) is full rank.

Geometrically, this means that the lines are all parallel. So we can say that \( A \) above is full rank, unless all the lines are parallel.

Here is another solution of the problem (or really, a variation on the solution given above). If we define

\[
C = \begin{bmatrix}
-v_1 & 0 & \cdots & 0 & I \\
0 & -v_2 & \cdots & 0 & I \\
\vdots & \vdots & \ddots & \vdots & I \\
0 & 0 & \cdots & -v_m & I
\end{bmatrix}, \quad d = \begin{bmatrix}
p_1 \\
\vdots \\
p_m
\end{bmatrix}, \quad u = \begin{bmatrix}
t_1 \\
\vdots \\
t_m \\
z
\end{bmatrix},
\]

we have

\[
\sum_{i=1}^m \text{dist}(z, \mathcal{L}_i)^2 = \min_{t_1, \ldots, t_m} \| C u - d \|,
\]

and

\[
\min_z \sum_{i=1}^m \text{dist}(z, \mathcal{L}_i)^2 = \min_u \| C u - d \|.
\]

In the last expression, we are optimizing over the line parameters \( t_i \) and the point \( z \) at the same time.

Therefore, assuming \( C \) is full rank, we have

\[
z^* = \begin{bmatrix}
0 & 0 \\
0 & I
\end{bmatrix} (C^T C)^{-1} C^T d,
\]

which expands to the same solution we have above. And of course, \( C \) is full rank if and only if \( A \) is, which occurs exactly when the lines are not all parallel.
(b) The following code solves for the point of closest convergence using the two different approaches and checks that the solutions are identical.

```matlab
% first solution
A=[];
b=[];
for i=1:m
    A=[A;eye(n)-V(:,i)*V(:,i)';]
    b=[b;(eye(n)-V(:,i)*V(:,i)')*P(:,i)];
end
zstar=A\b;

% second solution
C=zeros(n*m,m);
E=[];
d=[];
for i=1:m
    E=[E;eye(n)];
    C(n*(i-1)+1:n*i,i)=-V(:,i);
    d=[d;P(:,i)];
end
C=[C E];

zstar=A\b;
f=C\d;
zstar2=f(m+1:m+n);

% check that two solutions give (almost) same answer
zstar2-zstar
```

The result is $z^* = (1.9157, 3.3951)$ and figure 1 shows the lines together with the point of closest convergence.