

# HW2

January 18, 2021

## 1 CSE 252B: Computer Vision II, Winter 2021 – Assignment 2

1.0.1 Instructor: Ben Ochoa

1.0.2 Due: Wednesday, February 3, 2021, 11:59 PM

### 1.1 Instructions

- Review the academic integrity and collaboration policies on the course website.
- This assignment must be completed individually.
- This assignment contains both math and programming problems.
- All solutions must be written in this notebook
- Math must be done in Markdown/LaTeX.
- You must show your work and describe your solution.
- Programming aspects of this assignment must be completed using Python in this notebook.
- Your code should be well written with sufficient comments to understand, but there is no need to write extra markdown to describe your solution if it is not explicitly asked for.
- This notebook contains skeleton code, which should not be modified (This is important for standardization to facilitate efficient grading).
- You may use python packages for basic linear algebra, but you may not use packages that directly solve the problem. If you are uncertain about using a specific package, then please ask the instructional staff whether or not it is allowable.
- You must submit this notebook exported as a pdf. You must also submit this notebook as an .ipynb file.
- Your code and results should remain inline in the pdf (Do not move your code to an appendix).
- You must submit both files (.pdf and .ipynb) on Gradescope. You must mark each problem on Gradescope in the pdf.
- It is highly recommended that you begin working on this assignment early.

### 1.2 Problem 1 (Math): Line-plane intersection (5 points)

The line in 3D defined by the join of the points  $\mathbf{X}_1 = (X_1, Y_1, Z_1, T_1)^\top$  and  $\mathbf{X}_2 = (X_2, Y_2, Z_2, T_2)^\top$  can be represented as a Plucker matrix  $\mathbf{L} = \mathbf{X}_1\mathbf{X}_2^\top - \mathbf{X}_2\mathbf{X}_1^\top$  or pencil of points  $\mathbf{X}(\lambda) = \lambda\mathbf{X}_1 + (1 - \lambda)\mathbf{X}_2$  (i.e.,  $\mathbf{X}$  is a function of  $\lambda$ ). The line intersects the plane  $\boldsymbol{\pi} = (a, b, c, d)^\top$  at the point  $\mathbf{X}_L = \mathbf{L}\boldsymbol{\pi}$  or  $\mathbf{X}(\lambda_\pi)$ , where  $\lambda_\pi$  is determined such that  $\mathbf{X}(\lambda_\pi)^\top \boldsymbol{\pi} = 0$  (i.e.,  $\mathbf{X}(\lambda_\pi)$  is the point on  $\boldsymbol{\pi}$ ). Show that  $\mathbf{X}_L$  is equal to  $\mathbf{X}(\lambda_\pi)$  up to scale.

Your solution here

### 1.3 Problem 2 (Math): Line-quadric intersection (5 points)

In general, a line in 3D intersects a quadric  $Q$  at zero, one (if the line is tangent to the quadric), or two points. If the pencil of points  $\mathbf{X}(\lambda) = \lambda\mathbf{X}_1 + (1-\lambda)\mathbf{X}_2$  represents a line in 3D, the (up to two) real roots of the quadratic polynomial  $c_2\lambda^2 + c_1\lambda + c_0 = 0$  are used to solve for the intersection point(s)  $\mathbf{X}(\lambda_Q)$ . Show that  $c_2 = \mathbf{X}_1^\top Q \mathbf{X}_1 - 2\mathbf{X}_1^\top Q \mathbf{X}_2 + \mathbf{X}_2^\top Q \mathbf{X}_2$ ,  $c_1 = 2(\mathbf{X}_1^\top Q \mathbf{X}_2 - \mathbf{X}_2^\top Q \mathbf{X}_1)$ , and  $c_0 = \mathbf{X}_2^\top Q \mathbf{X}_2$ .

Your solution here

### 1.4 Problem 3 (Programming): Linear Estimation of the Camera Projection Matrix (15 points)

Download input data from the course website. The file `hw2_points3D.txt` contains the coordinates of 50 scene points in 3D (each line of the file gives the  $\tilde{X}_i$ ,  $\tilde{Y}_i$ , and  $\tilde{Z}_i$  inhomogeneous coordinates of a point). The file `hw2_points2D.txt` contains the coordinates of the 50 corresponding image points in 2D (each line of the file gives the  $\tilde{x}_i$  and  $\tilde{y}_i$  inhomogeneous coordinates of a point). The scene points have been randomly generated and projected to image points under a camera projection matrix (i.e.,  $\mathbf{x}_i = \mathbf{P}\mathbf{X}_i$ ), then noise has been added to the image point coordinates.

Estimate the camera projection matrix  $\mathbf{P}_{\text{DLT}}$  using the direct linear transformation (DLT) algorithm (with data normalization). You must express  $\mathbf{x}_i = \mathbf{P}\mathbf{X}_i$  as  $[\mathbf{x}_i]^\perp \mathbf{P}\mathbf{X}_i = \mathbf{0}$  (not  $\mathbf{x}_i \times \mathbf{P}\mathbf{X}_i = \mathbf{0}$ ), where  $[\mathbf{x}_i]^\perp \mathbf{x}_i = \mathbf{0}$ , when forming the solution. Return  $\mathbf{P}_{\text{DLT}}$ , scaled such that  $\|\mathbf{P}_{\text{DLT}}\|_{\text{Fro}} = 1$

The following helper functions may be useful in your DLT function implementation. You are welcome to add any additional helper functions.

```
[ ]: import numpy as np
import time

def Homogenize(x):
    # converts points from inhomogeneous to homogeneous coordinates
    return np.vstack((x,np.ones((1,x.shape[1])))

def Dehomogenize(x):
    # converts points from homogeneous to inhomogeneous coordinates
    return x[:-1]/x[-1]

def Normalize(pts):
    # data normalization of n dimensional pts
    #
    # Input:
    # pts - is in inhomogeneous coordinates
    # Outputs:
    # pts - data normalized points
    # T - corresponding transformation matrix
    """your code here"""
```

```

T = np.eye(pts.shape[0]+1)
return pts, T

def ComputeCost(P, x, X, scale=1):
    # Inputs:
    #   x - 2D inhomogeneous image points
    #   X - 3D inhomogeneous scene points
    #
    # Output:
    #   cost - Total reprojection error
    n = x.shape[1]
    covarx = np.eye(2*n)

    """your code here"""

    cost = np.inf
    return cost

```

```

[ ]: def DLT(x, X, normalize=True):
    # Inputs:
    #   x - 2D inhomogeneous image points
    #   X - 3D inhomogeneous scene points
    #   normalize - if True, apply data normalization to x and X
    #
    # Output:
    #   P - the (3x4) DLT estimate of the camera projection matrix
    P = np.eye(3,4)+np.random.randn(3,4)/10

    # data normalization
    if normalize:
        x, T = Normalize(x)
        X, U = Normalize(X)
    else:
        x = Homogenize(x)
        X = Homogenize(X)

    """your code here"""

    # data denormalize
    if normalize:

```

```

        P = np.linalg.inv(T) @ P @ U

    return P

def displayResults(P, x, X, title):
    print(title+' =')
    print (P/np.linalg.norm(P)*np.sign(P[-1,-1]))

# load the data
x=np.loadtxt('hw2_points2D.txt').T
X=np.loadtxt('hw2_points3D.txt').T

# compute the linear estimate without data normalization
print ('Running DLT without data normalization')
time_start=time.time()
P_DLT = DLT(x, X, normalize=False)
cost = ComputeCost(P_DLT, x, X)
time_total=time.time()-time_start
# display the results
print('took %f secs'%time_total)
print('Cost=%.9f'%cost)

# compute the linear estimate with data normalization
print ('Running DLT with data normalization')
time_start=time.time()
P_DLT = DLT(x, X, normalize=True)
cost = ComputeCost(P_DLT, x, X)
time_total=time.time()-time_start
# display the results
print('took %f secs'%time_total)
print('Cost=%.9f'%cost)

print("\n==Correct outputs==")
print("Cost=%.9f without data normalization"%97.053718991)
print("Cost=%.9f with data normalization"%84.104680130)

```

```

[ ]: # Report your P_DLT value here!
displayResults(P_DLT, x, X, 'P_DLT')

```

### 1.5 Problem 4 (Programming): Nonlinear Estimation of the Camera Projection Matrix (30 points)

Use  $P_{DLT}$  as an initial estimate to an iterative estimation method, specifically the Levenberg-Marquardt algorithm, to determine the Maximum Likelihood estimate of the camera projection matrix that minimizes the projection error. You must parameterize the camera projection matrix

as a parameterization of the homogeneous vector  $\mathbf{p} = \text{vec}(\mathbf{P}^\top)$ . It is highly recommended to implement a parameterization of homogeneous vector method where the homogeneous vector is of arbitrary length, as this will be used in following assignments.

Report the initial cost (i.e. cost at iteration 0) and the cost at the end of each successive iteration. Show the numerical values for the final estimate of the camera projection matrix  $\mathbf{P}_{\text{LM}}$ , scaled such that  $\|\mathbf{P}_{\text{LM}}\|_{\text{Fro}} = 1$ .

The following helper functions may be useful in your LM function implementation. You are welcome and encouraged to add any additional helper functions.

Hint: LM has its biggest cost reduction after the 1st iteration. You'll know if you are implementing LM correctly if you experience this.

```
[ ]: # Note that np.sinc is different than defined in class
def Sinc(x):
    # Returns a scalar valued sinc value
    """your code here"""

    y = x
    return y

def Jacobian(P,p,X):
    # compute the jacobian matrix
    #
    # Input:
    # P - 3x4 projection matrix
    # p - 11x1 homogeneous parameterization of P
    # X - 3n 3D scene points
    # Output:
    # J - 2nx11 jacobian matrix
    J = np.zeros((2*X.shape[1],11))

    """your code here"""

    return J

def Parameterize(P):
    # wrapper function to interface with LM
    # takes all optimization variables and parameterizes all of them
    # in this case it is just P, but in future assignments it will
    # be more useful
    return ParameterizeHomog(P.reshape(-1,1))
```

```

def Deparameterize(p):
    # Deparameterize all optimization variables
    return DeParameterizeHomog(p).reshape(3,4)

def ParameterizeHomog(V):
    # Given a homogeneous vector V return its minimal parameterization
    """your code here"""

    return v_hat

def DeParameterizeHomog(v):
    # Given a parameterized homogeneous vector return its deparameterization
    """your code here"""

    return v_bar

```

```

[ ]: def LM(P, x, X, max_iters, lam):
    # Input:
    #   P - initial estimate of P
    #   x - 2D inhomogeneous image points
    #   X - 3D inhomogeneous scene points
    #   max_iters - maximum number of iterations
    #   lam - lambda parameter
    # Output:
    #   P - Final P (3x4) obtained after convergence

    # data normalization
    x, T = Normalize(x)
    X, U = Normalize(X)

    """your code here"""

    # you may modify this so long as the cost is computed
    # at each iteration
    for i in range(max_iters):

```

```

        cost = ComputeCost(P, x, X)
        print ('iter %03d Cost %.9f'%(i+1, cost))

    # data denormalization
    P = np.linalg.inv(T) @ P @ U
    return P

# LM hyperparameters
lam = .001
max_iters = 100

# Run LM initialized by DLT estimate with data normalization
print ('Running LM with data normalization')
print ('iter %03d Cost %.9f'%(0, cost))
time_start=time.time()
P_LM = LM(P_DLT, x, X, max_iters, lam)
time_total=time.time()-time_start
print('took %f secs'%time_total)

print("\n==Correct outputs==")
print("Begins at %.9f; ends at %.9f"%(84.104680130, 82.790238005))

```

```

[ ]: # Report your P_LM final value here!
displayResults(P_LM, x, X, 'P_LM')

```