Tasks for lecture 4

Recall the least squares problems Pontius and Filip from lecture 2 and 3. We continue the work on these here.

- Solve the exercises from the lecture slides.
- Solve the problems of both datasets with SVD and print relevant information (continuation from lecture 3).
- Perform error estimations of your solutions (create different thresholds than the default and *SVD::eps*).

Exercises lecture 4 (+ Mandatory 1 hint)

Design matrix and right hand side vector

- Choose sigma (std)

$$\mathbf{A}_{ij} \;:=\; \mathbf{A}_{ij}/\sigma_i \quad i = 1,\ldots m, \quad j = 1,\ldots n$$

$$\mathbf{b}_i \;:=\; \mathbf{b}_i/\sigma_i \quad i = 1,\ldots m$$

$$\boxed{\text{The SVD solution } \mathbf{x} = \mathbf{V}\tilde{\mathbf{W}}^{-1}\mathbf{U}^T\mathbf{b}}$$

Residual error (std of residuals) < Random fitting

$$\boxed{\epsilon_{residual} = \frac{\|\mathbf{A}\mathbf{x} - \mathbf{b}\|}{\|\mathbf{b}\|}} \qquad \epsilon_{residual} \simeq \sqrt{\frac{m-n}{m}}$$

Error on parameter estimates (std)

$$\boxed{[\delta\mathbf{x}]_j \simeq \sqrt{\sum_{i=1}^{n}\left(\frac{V_{ji}}{w_i}\right)^2} \quad j = 1,\ldots,n}$$

Certified Regression Statistics

Parameter	Estimate	Standard Deviation of Estimate
B0	-1467.48961422980	298.084530995537
B1	-2772.17959193342	559.779865474950
B2	-2316.37108160893	466.477572127796
B3	-1127.97394098372	227.204274477751
B4	-354.478233703349	71.6478660875927
B5	-75.1242017393757	15.2897178747400
B6	-10.8753180355343	2.23691159816033
B7	-1.06221498588947	0.221624321934227
B8	-0.670191154593408E-01	0.142363763154724E-01
B9	-0.246781078275479E-02	0.535617408889821E-03
B10	-0.402962525080404E-04	0.896632837373868E-05

Residual
Standard Deviation 0.334801051324544E-02

University of
Southern Denmark

Exercises lecture 4

Ex. 2.1 Check if the following vectors are linearly independent

i) $\mathbf{x}_1 = (1,1,0)$; $\mathbf{x}_2 = (0,1,1)$; $\mathbf{x}_3 = (1,2,1)$

ii) $\mathbf{x}_1 = (1,1,0,0)$; $\mathbf{x}_2 = (0,1,1,0)$; $\mathbf{x}_3 = (0,0,1,1)$

iii) $\mathbf{x}_1 = (1,1,8)$; $\mathbf{x}_2 = (8,1,-5)$; $\mathbf{x}_3 = (0,0,0)$

iv) $\mathbf{x}_1 = (1,1,8,2,4)$; $\mathbf{x}_2 = (8,1,-5,3,2)$; $\mathbf{x}_3 = (4,5,1,-2,3)$; $\mathbf{x}_4 = (2,7,-4,3,8)$; $\mathbf{x}_5 = (-4,9,2,-21,-8)$

I. x_3 Linear combination of x_1 and x_2 → Linear dependent
II. No linear combination possible → Linear independent
III. Definition: Null vector in the set → Linear dependent
IV. Reduced row echelon form shows linear combination resulting in null vector → Linear dependent

I. x_3 Linear combination of x_1 and x_2 → Linear dependent
II. No linear combination possible → Linear independent
III. Definition: Null vector in the set → Linear dependent
IV. Reduced row echelon form shows linear combination resulting in null vector → Linear dependent

Exercises lecture 4

Ex. 2.2 Assume the columns of a 3×3 matrix $\mathbf{A}$ is given by the vectors in Exercise 2.1 ii). Compute $\mathbf{A}^T\mathbf{A}$.

```
>> A = [1,1,0;0,1,0;0,0,1]

A =

     1     1     0
     0     1     0
     0     0     1

>> transpose(A)*A

ans =

     1     1     0
     1     2     0
     0     0     1
```

University of
Southern Denmark

Exercises lecture 4

Ex. 3.1 Check if the following vectors are orthogonal, respectively orthonormal

i) $\mathbf{x}_1 = (1, 1, 0, 0);\quad \mathbf{x}_2 = (0, 0, 1, 1);$

ii) $\mathbf{x}_1 = (1, 1, 0);\quad \mathbf{x}_2 = (0, 1, 1);$

iii) $\mathbf{x}_1 = (\frac{\sqrt{3}}{2}, \frac{1}{2});\quad \mathbf{x}_2 = (-\frac{1}{2}, \frac{\sqrt{3}}{2});$

iv) $\mathbf{x}_1 = (1, 3, 8);\quad \mathbf{x}_2 = (0, 0, 0);$

Use definition 5

Dot product between vectors

	Orthogonal	Orthonormal
i	Y	N
ii	N	N
iii	Y	Y
iv	Y	N

Exercises lecture 4

Ex. 3.2 Consider a 3×3 rotation matrix $\mathbf{R}$. Is this matrix always orthonormal ?

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

Definition 6 and 7 → Columns always give a dot product of 1

Exercises lecture 4

Ex. 6.1 Write a computer program (in any computer language you like) that implements the Gram Schmidt method.

```
function E = Gram_Schmidt(X)
    [n,k] = size(X);
    E = zeros(n,k);
    E(:,1) = X(:,1)/ norm(X(:,1));
    for i=2:k
        E(:,i) = X(:,i);
        for j=1:i-1
            E(:,i)=E(:,i)-dot(X(:,i),E(:,j))*E(:,j);
        end
        length =  norm(E(:,i));
        %disp(length);
        if length > 10^(-15)
            E(:,i) = E(:,i) / norm(E(:,i));
        else
            E(:,i) = zeros(n,1);
        end

    end
    disp("E = ");
    disp(E);
end
```

Gram-Schmidt program

Set of linear independent vectors transformed to a set of orthonormal vectors that span the same space

The Gram-Schmidt method:

$\mathbf{e}_1 := \mathbf{x}_1/\|\mathbf{x}_1\|$

For $i := 2, \ldots, k$ do {

$\mathbf{e}_i := \mathbf{x}_i - \sum_{j=1}^{i-1}(\mathbf{x}_i \cdot \mathbf{e}_j)\mathbf{e}_j$

$\mathbf{e}_i := \mathbf{e}_i/\|\mathbf{e}_i\|$

}

Exercises lecture 4

Ex. 6.2 Without proof, we can assume that the vectors $\mathbf{x}_1 = (2, 8, 4, 2, 1)$; $\mathbf{x}_2 = (1, 1, 5, 7, 8)$; $\mathbf{x}_3 = (4, -5, 1, -4, 3)$ are linearly independent vectors in $\mathbb{R}^5$. Use the program from Ex. 6.1 to compute an orthonormal basis for the subspace of $\mathbb{R}^5$ spanned by $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$.

```
function E = Gram_Schmidt(X)
    [n,k] = size(X);
    E = zeros(n,k);
    E(:,1) = X(:,1)/ norm(X(:,1));
    for i=2:k
        E(:,i) = X(:,i);
        for j=1:i-1
            E(:,i)=E(:,i)-dot(X(:,i),E(:,j))*E(:,j);
        end
        length =  norm(E(:,i));
        %disp(length);
        if length > 10^(-15)
            E(:,i) = E(:,i) / norm(E(:,i));
        else
            E(:,i) = zeros(n,1);
        end

    end
    disp("E = ");
    disp(E);
end
```

```
x1 =

     2
     8
     4
     2
     1


x2 =

     1
     1
     5
     7
     8


x3 =

     4
    -5
     1
    -4
     3

E =
    0.2120   -0.0161    0.6657
    0.8480   -0.3509   -0.1936
    0.4240    0.2543    0.2811
    0.2120    0.5570   -0.5977
    0.1060    0.7083    0.2883
```

Exercises lecture 4

I. Trivial null space due to linear independency $\mathbf{A} \cdot \mathbf{x} = 0.$

II. Range and spanned vector space (rank)

```
Nullspace N(A)  Matrix 3x0:




Nullity of A: 0
Range B(A)      Matrix 4x3:
        0.184815        0.861674        0.335038
        0.567225       -0.401294         0.71918
        0.611814        0.258264       -0.338436
       -0.519407        0.172573        0.505955

Rank of A: 3
```

Ex. 6.5 Consider the matrix $\mathbf{A}$ given as

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 2 \\ 0 & 4 & 4 \\ 2 & 4 & 4 \\ -1 & -4 & -3 \end{pmatrix}$$

i) Determine the Null Space $N(\mathbf{A})$ and an orthonormal basis for $N(\mathbf{A})$

ii) Determine the range $B(\mathbf{A})$ and an orthonormal basis for $B(\mathbf{A})$

What has all this to do with singular value decomposition? SVD explicitly constructs orthonormal bases for the nullspace and range of a matrix! Specifically, the columns of $\mathbf{U}$ whose same-numbered elements w_j are *nonzero* are an orthonormal set of basis vectors that span the range; the columns of $\mathbf{V}$ whose same-numbered elements w_j are *zero* are an orthonormal basis for the nullspace. Our SVD object has methods that return the rank or nullity (integers), and also the range and nullspace, each of these packaged as a matrix whose columns form an orthonormal basis for the respective subspace.

Page 67-69 Chapter 2.6.1

SVD YT videos of Steve Brunton databookuw.com/

Exercises lecture 4

```
function b_ls = least_square(E, b)
    k= size(E,2);
    b_ls = zeros(size(E,1),1);
    for i=1:k
        b_ls = b_ls + dot(b,E(:,i))*E(:,i)
    end
end
```

Theorem 5 Let $\mathbf{u}_1 \ldots, \mathbf{u}_K$ be an arbitrary orthonormal basis for $B(\mathbf{A})$. Then the least squares solution $\mathbf{x}$ that minimizes $\|\mathbf{Ax} - \mathbf{b}\|$ satisfies

$$\mathbf{Ax} = \sum_{k=1}^{K} (\mathbf{b} \cdot \mathbf{u}_k)\mathbf{u}_k \equiv \mathbf{b}_{LS}$$

Ex. 6.3 Compute the point in the subspace spanned by $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ from Ex. 6.2 that is nearest to $(5, 6, 1, 2, 3)$ (HINT: Consider the statement in Theorem 5 in the next section and use your result from Ex. 6.2).

```
x1 = [2;8;4;2;1]
x2 = [1;1;5;7;8]
x3 = [4;-5;1;-4;3]
X = [x1,x2,x3];

E = Gram_Schmidt(X);


%6.3

b =  [5;6;1;2;3];
b_ls = least_square(E,b);
```

```
b_ls =

    2.9391
    5.3336
    4.0288
    1.0129
    2.3116
```

University of Southern Denmark