EECS 280 Computer Vision
Spring 2025 Jitendra Malik, Angjoo Kanazawa

This homework is due on Monday, Jan 27, 2025, at 11:59PM.

1 Pose, Shape and Geometric Transformations

Points on an object can be characterized by their 3D coordinates with respect to the camera coordinate system. But what happens, when we move the object? In a certain sense when a chair is moved in 3D space, it remains the "same" even though the coordinates of points on it with respect to the camera (or any fixed) coordinate system do change. This distinction is captured by the use of the terms **pose** and **shape**.

- *Pose:* The position and orientation of the object with respect to the camera. This is specified by 6 numbers (3 for its translation, 3 for rotation). For example, we might consider the coordinates of the centroid of the object relative to the center of projection, and the rotation of a coordinate frame on the object with respect to that of the camera.
- *Shape:* The coordinates of the points of an object relative to a coordinate frame on the object. These remain invariant when the object undergoes rotations and translations.

To make these notions more precise, we need to develop the basic theory of **Euclidean Transformations**. The set of transformations defines a notion of "congruence" or having the same shape. In high school geometry we learned that two planar triangles are congruent if one of them can be rotated and translated so as to lie exactly on top of another. Rotation and translation are examples of Euclidean transformations, also known as **isometries** or **rigid body motions**, defined as transformation that preserve distances between any pair of points. When I move a chair, this holds true between any pair of points on the chair, but obviously not for points on a balloon that is being inflated.

In this chapter we will review the basic concepts relevant to Euclidean transformations. Then we will study a more general class of transformations, called **affine transformations**, which include Euclidean transformations as a subset. The set of **projective transformations** is even more general, and is a superset of affine transformations. All three classes of transformations find utility in a study of vision.

2 Euclidean Transformations

$\mathbf{A}$	Matrix
$\mathbf{a}$	Vector
$\mathbf{I}$	The identity matrix
$\psi : \mathbb{R}^n \mapsto \mathbb{R}^n$	Transformation
$\mathbf{x} \cdot \mathbf{y}$	Dot product (scalar product)
$\mathbf{x} \wedge \mathbf{y}$	Cross product (vector product)
$\|\|\mathbf{x}\|\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$	Norm

Definition 1 *Euclidean transformations (also known as isometries) are transformations that preserve distances between pairs of points.*

$$||\psi(\mathbf{a}) - \psi(\mathbf{b})|| = ||\mathbf{a} - \mathbf{b}|| \tag{1}$$

Translations, $\psi(\mathbf{a}) = \mathbf{a} + \mathbf{t}$, are isometries, since

$$||\psi(\mathbf{a}) - \psi(\mathbf{b})|| = ||\mathbf{t} + \mathbf{a} - (\mathbf{t} + \mathbf{b})|| = ||\mathbf{a} - \mathbf{b}|| \tag{2}$$

We now define orthogonal transformations; these constitute another major class of isometries.

Definition 2 *A linear transformation:* $\psi(\mathbf{a}) = \mathbf{A}\mathbf{a}$*, for some matrix* $\mathbf{A}$*.*

Definition 3 *Orthogonal transformations are linear transformations which preserve inner products.*

$$\mathbf{a} \cdot \mathbf{b} = \psi(\mathbf{a}) \cdot \psi(\mathbf{b}) \tag{3}$$

Property 1 *Orthogonal transformations preserve norms.*

$$\mathbf{a} \cdot \mathbf{a} = \psi(\mathbf{a}) \cdot \psi(\mathbf{a}) \implies ||\mathbf{a}|| = ||\psi(\mathbf{a})|| \tag{4}$$

Property 2 *Orthogonal transformations are isometries.*

$$(\psi(\mathbf{a}) - \psi(\mathbf{b})) \cdot (\psi(\mathbf{a}) - \psi(\mathbf{b})) \stackrel{?}{=} (\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) \tag{5}$$

$$||\psi(\mathbf{a})||^2 + ||\psi(\mathbf{b})||^2 - 2(\psi(\mathbf{a}) \cdot \psi(\mathbf{b})) \stackrel{?}{=} ||\mathbf{a}||^2 + ||\mathbf{b}||^2 - 2(\mathbf{a} \cdot \mathbf{b}) \tag{6}$$

By property 1,

$$||\psi(\mathbf{a})||^2 = ||\mathbf{a}||^2 \tag{7}$$

$$||\psi(\mathbf{b})||^2 = ||\mathbf{b}||^2. \tag{8}$$

By definition 3,

$$\psi(\mathbf{a}) \cdot \psi(\mathbf{b}) = \mathbf{a} \cdot \mathbf{b}. \tag{9}$$

Thus, equality holds.

Note that translations do not preserve norms (the distance with respect to the origin changes) and are not even linear transformations, except for the trivial case of translation by $\mathbf{0}$.

2.1 Properties of orthogonal matrices

Let ψ be an orthogonal transformation whose action we can represent by matrix multiplication, $\psi(\mathbf{a}) = \mathbf{A}\mathbf{a}$. Then, because it preserves inner products:

$$\psi(\mathbf{a}) \cdot \psi(\mathbf{b}) = \mathbf{a}^T\mathbf{b}\,. \tag{10}$$

By substitution,

$$\psi(\mathbf{a}) \cdot \psi(\mathbf{b}) = (\mathbf{A}\mathbf{a})^T(\mathbf{A}\mathbf{b}) \tag{11}$$

$$= \mathbf{a}^T\mathbf{A}^T\mathbf{A}\mathbf{b}\,. \tag{12}$$

Thus,

$$\mathbf{a}^T\mathbf{b} = \mathbf{a}^T\mathbf{A}^T\mathbf{A}\mathbf{b} \implies \mathbf{A}^T\mathbf{A} = \mathbf{I} \implies \mathbf{A}^T = \mathbf{A}^{-1}\,. \tag{13}$$

Note that $\det(\mathbf{A})^2 = 1$ which implies that $\det(\mathbf{A}) = +1$ or -1. Each column of $\mathbf{A}$ has norm 1, and is orthogonal to the other column.

In 2D, these constraints put together force $\mathbf{A}$ to be one of two types of matrices.

$$\underbrace{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}_{\text{rotation, det}=+1} \text{or} \underbrace{\begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}}_{\text{reflection, det}=-1}$$

Under a rotation by angle θ,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \text{and} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

The reflection matrix above corresponds to reflection around the line with angle $\frac{\theta}{2}$ (verify). Note that two rotations one after the other give another rotation, while two reflections give us a rotation.

Let us now construct some examples in 3D. Just as in 2D, rotations are characterized by orthogonal matrices with $\det = +1$. For orthogonal matrices, each column vector has length 1, and the dot product of any two different columns is 0. This gives rise to six constraints (3 pairwise dot product constraints, and 3 length constraints), so for a 3 dimensional rotation matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{14}$$

with 9 total parameters, there are really only three free parameters. There are several methods by which these parameters can be specified, as we will study later. Here are a few example rotation matrices.

- Rotation about z-axis by θ:

$$\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{15}$$

- Rotation about x-axis by θ:

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \tag{16}$$

2.2 Group structure of isometries

Any isometry can be expressed as the combination of an orthogonal transformation followed by a translation as follows:

$$\psi(\mathbf{a}) = \mathbf{A}\mathbf{a} + \mathbf{t} \tag{17}$$

where $\mathbf{A}$ represents the orthogonal matrix and $\mathbf{t}$ is the translation vector.

The set of rigid body motions constitutes a *group*[1]. In our notation, $\psi_1 \circ \psi_2$, ψ_1 composed with ψ_2, denotes that we apply ψ_2 first and then ψ_1.

[1]A group $(G, \circ)$ is a set G with a binary operation $\circ$ that satisfies the following four axioms: Closure: For all a, b in G, the result of $a \circ b$ is also in G. Associativity: For all a, b and c in G, $(a \circ b) \circ c = a \circ (b \circ c)$. Identity element: There exists an element e in G such that for all a in G, $e \circ a = a \circ e = a$. Inverse element: For each a in G, there exists an element b in G such that $a \circ b = b \circ a = e$, where e is an identity element.

We will show first that isometries are closed under composition. Consider two rigid body motions, ψ_1 and ψ_2:

$$\psi_1(\mathbf{a}) = \mathbf{A}_1\mathbf{a} + \mathbf{t}_1 \qquad \psi_2(\mathbf{a}) = \mathbf{A}_2\mathbf{a} + \mathbf{t}_2. \tag{18}$$

Then we have

$$\begin{aligned}
\psi_1 \circ \psi_2(\mathbf{a}) &= \mathbf{A}_1(\mathbf{A}_2\mathbf{a} + \mathbf{t}_2) + \mathbf{t}_1 && (19)\\
&= \mathbf{A}_1\mathbf{A}_2\mathbf{a} + \mathbf{A}_1\mathbf{t}_2 + \mathbf{t}_1 && (20)\\
&= (\mathbf{A}_1\mathbf{A}_2)\mathbf{a} + (\mathbf{A}_1\mathbf{t}_2 + \mathbf{t}_1) && (21)\\
&= \mathbf{A}_3\mathbf{a} + \mathbf{t}_3 && (22)
\end{aligned}$$

where $\mathbf{A}_3 = \mathbf{A}_1\mathbf{A}_2$ and $\mathbf{t}_3 = \mathbf{A}_1\mathbf{t}_2 + \mathbf{t}_3$. Thus, $\psi_1 \circ \psi_2 = \psi_3$ is also a rigid body motion, under the assumption that the product of two orthogonal matrices is orthogonal (Verify!)

Note that translations and rotations are closed under composition, but reflections are not.

We can verify the remaining axioms for showing that isometries constitute a group

- Identity: $\mathbf{A} = \mathbf{I}, \mathbf{d} = 0$.
- Inverse: We need $\mathbf{A}_1\mathbf{A}_2 = \mathbf{I}$ and $\mathbf{t}_3 = \mathbf{A}_1\mathbf{t}_2 + \mathbf{t}_1 = 0$. This means that for ψ_1 to be the inverse of ψ_2, $\mathbf{A}_1 = \mathbf{A}_2^T$ and $\mathbf{d}_2 = -\mathbf{A}_1^{-1}\mathbf{t}_1$
- Associativity: left as an exercise for the reader.

3 Parametrizing Rotations in 3D

Recall that rotation matrices have the property that each column vector has length 1 and the dot product of any 2 different columns is 0. These 6 constraints leave only 3 degrees of freedom. Here are some alternative notations used to represent orthogonal matrices in 3-D:

- Euler angles which specify rotations about 3 axes
- Axis plus amount of rotation
- Quaternions which generalize complex numbers from 2-D to 3-D. (Note, a complex number can represent a rotation in 2-D)

We will use the axis and rotation as the preferred representation of an orthogonal matrix: $\mathbf{s}$, θ, where $\mathbf{s}$ is the unit vector of the axis of rotation and θ is the amount of rotation.

Definition 4 *A matrix* $\mathbf{S}$ *is skew-symmetric if* $\mathbf{S} = -\mathbf{S}^T$.

Skew symmetric matrices can be used to represent "cross" products or vector products. Recall:

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \wedge \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{bmatrix}$$

We define $\widehat{\mathbf{a}}$ as:

$$\widehat{\mathbf{a}} \stackrel{def}{=} \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

Thus, multiplying $\widehat{\mathbf{a}}$ by any vector gives:

$$\begin{aligned} \widehat{\mathbf{a}} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} &= \begin{bmatrix} -a_3 b_2 + a_2 b_3 \\ a_3 b_1 - a_1 b_3 \\ -a_2 b_1 + a_1 b_2 \end{bmatrix} \\ &= \mathbf{a} \wedge \mathbf{b} \end{aligned}$$

Consider now, the equation of motion of a point q on a rotating body:

$$\dot{\mathbf{q}}(t) = \omega \wedge \mathbf{q}(t)$$

where the direction of ω specifies the axis of rotation and $\|\omega\|$ specifies the angular speed. Rewriting with $\widehat{\omega}$

$$\dot{\mathbf{q}}(t) = \widehat{\omega}\mathbf{q}(t)$$

The solution of this differential equation involves the exponential of a matrix. (In matlab, this is the operator `expm`.)

$$\mathbf{q}(t) = e^{\widehat{\omega}t}\mathbf{q}(0)$$

Where,

$$e^{\widehat{\omega}t} = \mathbf{I} + \widehat{\omega}t + \frac{(\widehat{\omega}t)^2}{2!} + \frac{(\widehat{\omega}t)^3}{3!} + ...$$

Collecting the odd and even terms in the above equation, we get to **Roderigues Formula** for a rotation matrix $\mathbf{R}$.

$$\begin{aligned} \mathbf{R} &= e^{\phi\widehat{s}} \\ &= \mathbf{I} + \sin\phi\,\widehat{\mathbf{s}} + (1 - \cos\phi)\widehat{\mathbf{s}}^2 \end{aligned}$$

Here $\mathbf{s}$ is a unit vector along ω and $\phi = \|\omega\|t$ is the total amount of rotation. Given an axis of rotation, $\mathbf{s}$, and amount of rotation ϕ we can construct $\widehat{\mathbf{s}}$ and plug it in.

4 Affine transformations

Thus far we have focused on Euclidean transformations, $\psi(\mathbf{a}) = \mathbf{A}\mathbf{a} + \mathbf{t}$, where $\mathbf{A}$ is an orthogonal matrix. If we allow $\mathbf{A}$ to be any non-singular matrix (i.e., $\det \mathbf{A} \neq 0$), then we get the set of affine transformations. Note that the Euclidean transformations are a subset of the affine transformations.

4.1 Degrees of freedom

Let us count the degrees of freedom in the parameters that specify a transformation. For $\psi : \mathbb{R}^2 \mapsto \mathbb{R}^2$, Euclidean transformations have 3 free parameters (1 rotation, 2 translation), whereas Affine transformations have 6 (4 in $\mathbf{A}$ and 2 in $\mathbf{t}$). For $\psi : \mathbb{R}^3 \mapsto \mathbb{R}^3$, Euclidean transformations have 6 free parameters (3 rotation, 3 translation), whereas Affine transformations have 12 (9 in $\mathbf{A}$ and 3 in $\mathbf{t}$).

5 Written Exercises

(a) Show that in $\mathcal{R}^2$ reflection about the $\theta = \alpha$ line followed by reflection about the $\theta = \beta$ is equivalent to a rotation of $2(\beta - \alpha)$.

(b) Verify Roderigues formula by considering the powers of the skew-symmetric matrix associated with the cross product with a vector.

(c) Write a Python function for computing the orthogonal matrix $\mathbf{R}$ corresponding to rotation ϕ about the axis vector $\mathbf{s}$. Find the eigenvalues and eigenvectors of the orthogonal matrices and study any relationship to the axis vector. Verify the formula $\cos\phi = \frac{1}{2}\{\text{trace}(\mathbf{R}) - 1\}$. Show some points before and after the rotation has been applied.

(d) Write a Python function for the converse of that in the previous problem i.e. given an orthogonal matrix $\mathbf{R}$, compute the axis of rotation $\mathbf{s}$ and ϕ). Hint: Show that $\mathbf{R} - \mathbf{R}^T = (2\sin\phi)\widehat{\mathbf{s}}$

(e) Use least squares to find the best estimate of the Euclidean planar transformation (translation + rotation) E that minimizes the error $\sum_{j=1}^{4} \mid E\mathbf{u}_j - \mathbf{v}_j \mid^2$. Here

$$\mathbf{u}_j = [(-3,0),(1,1),(1,0),(1,-1)]$$

and

$$\mathbf{v}_j = [(0,3),(1,0),(0,0),(-1,0)]$$

(f) Show that the vanishing points of lines on a plane lie on the vanishing line of the plane.

6 Coding Exercise: AutoGrad

Fill out and turn in the following colab notebook(TODO:1,2,3,4,5) **with the cell outputs displayed**. You may use Andrej Karpathy's micrograd tutorial for reference. Additionally, answer the questions below.

(a) Are we strictly required to use our atomic operations when defining new functions, e.g., sigmoid? Under what conditions can we define new operations?

(b) When performing backpropagation on a Value, why do we accumulate the gradient as opposed to directly assigning the gradient?

7 Submission

Submit a single PDF containing written solutions, as well as the notebook (also as a PDF) at the very end. Submit via Gradescope.