1 Numerical Exercises

1. In the lecture you have seen the Jacobian which contains the derivatives of a vector-valued function w.r.t its variables. The cases presented in the lecture correspond to projective transformations (e.g. straight lines remain straight). However, for certain applications (e.g. cartography) the transformations are non-projective.

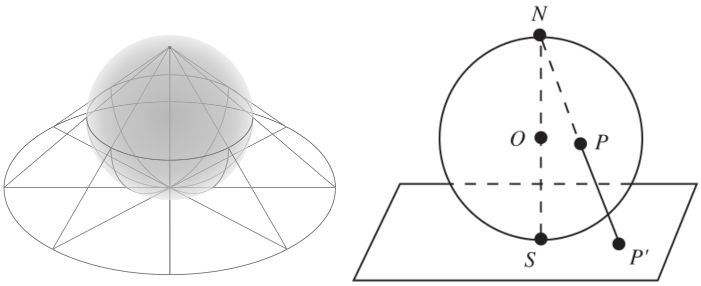

Figure 1: Illustration of the stereographic projection

The most prominent non-projective projection (i.e. straight lines do not remain straight) is the *stereographic projection*[1] which has already been used by the ancient greeks for celestial charts! This projection maps points on a sphere to a planar map. Hence it is widely used in catography.

Given a sphere of radius R and a set of points (x_s, y_s, z_s) on the sphere, the corresponding points on a planar map $(x_p, y_p) = f(x_s, y_s, z_s)$ are given by

$$x_p = \frac{x_s}{1 - z_s}$$
$$y_p = \frac{y_s}{1 - z_s}$$

Derive the Jacobian J of the projection function $f : \mathbb{R}^3 \mapsto \mathbb{R}^2$ w.r.t. the vector $\mathbf{x_s} = [x_s, y_s, z_s]$.

Solution
The Jacobian has to be of size 2×3 as there is two variables x_p and y_p that depend on three variables. Taking the derivative of x_p w.r.t. x_s leads to

$$\frac{\mathrm{d}x_p}{\mathrm{d}x_s} = \frac{-1}{z - 1}$$

which is the top left element of the jacobian. The other elements are computed similarly. The final jacobian is given by

$$J = \begin{bmatrix} -1/(z-1), & 0 & x/(z-1)^2 \\ 0 & -1/(z-1) & y/(z-1)^2 \end{bmatrix}$$

2. Given a general bijective transformation (not neccessarily a projective transformation), what can we say about its Jacobian?
Hint: think about: rank, determinant, invertibility, etc.

1 **Solution**
We consider continuously differentiable transformations as this type is typcially encountered in computer vision.

A bijective function is invertible and this property also holds for the jacobian of the function. Intuitively, if the function is bijective it can have zero gradient only at one point because otherwise this would conflict with bijectiveness. If the Jabocian is invertible, it is guaranteed to have full rank and a non-zero determinant.

[1] `https://en.wikipedia.org/wiki/Stereographic_projection`

If the determinant is positive, the transformation preserves orientation whereas a negative determinant corresponds to a reversal of orientation. Magnitude of the determinant describes how much an area is scaled (det = 1 is no scaling). Looking at the determinant from the previous subquestion we can see that the stereographic projection does not preserve area. You all probably know that intuitively because Antarctica (on a map) is overly large.

3. The KLT-Tracker uses a Gauss-Newton Update step as shown in the lecture and we will study the properties of this optimization method in this exercise.

   In general, the Gauss-Newton method solves sum-of-least squares optimization problems of the form
$$\min_{x\in\mathbb{R}^n} \frac{1}{2}\sum_{i=1}^{m}(f_i(x))^2 \Longleftrightarrow \min_{x\in\mathbb{R}^n} \frac{1}{2}||f(x)||^2$$
   For a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ and a step-size λ an update step is given by
$$x_{n+1} = x_n - \lambda \left[J(x_n)^\top J(x_n)\right]^{-1} J(x_n)^\top f(x_n)$$
   where J denotes the Jacobian of $f(x)$.

   Derive the above update step.
   Hint: Use a first-order taylor expansion around the current point x_n.

   Solution Let $g(x)$ denote the first-order taylor expansion of $f(x)$ around the current x_n.
$$g(x) = f(x_n) + \nabla f(x_n)^\top (x - x_n)$$
   With this, we can rewrite the optimization problem as
$$\min_{x\in\mathbb{R}^n} \frac{1}{2}||g(x)||^2 = \min_{x\in\mathbb{R}^n} \frac{1}{2}||J(x - x_n) + f(x_n)||^2$$
   Note we consider the Jacobian J at the point x_n.

   Let's look at the gradient of the function to be optimized.
$$\nabla\frac{1}{2}||g(x)||^2 = J^\top \left(J(x - x_n) + f(x_n)\right)$$
   Setting the gradient to zero (optimum!) and the solving for the optimal x yields
$$x = x_n - (J^\top J)^{-1} J^\top f(x_n)$$
   By introducing a step-size α to dampen the update step, the above expression directly leads to the statement given in the question
$$x_{n+1} = x_n - \lambda \left[J(x_n)^\top J(x_n)\right]^{-1} J(x_n)^\top f(x_n)$$

4. Similar to before, we consider a Gauss-Newton update step. Show that the method has second-order convergence in cases where the function $f(x)$ is nearly linear.
   Hint: Use a second-order taylor expansion and consider the Hessian at the current point x_n.
   Hint: What implications does 'nearly linear' have for the Hessian matrix?

   Solution
   If we consider a second-order taylor expansion around the point x_n, we get
$$h(x) = f(x_n) + \nabla f(x_n)^\top (x - x_n) + \frac{1}{2}(x - x_n)^\top H(x_n)(x - x_n)$$
   where $H(x_n)$ is the hessian matrix of the function f evaluated at point x_n.

   For functions that are nearly linear, the hessian matrix is close to zero, such that the second order term has negligible influence. In such a case the Gauss-Newton update step is similar to a Newton-Methods update step. Newtons method has second order convergence and hence for functions with small hessians, the Gauss-Newton method also has quadratic convergence.