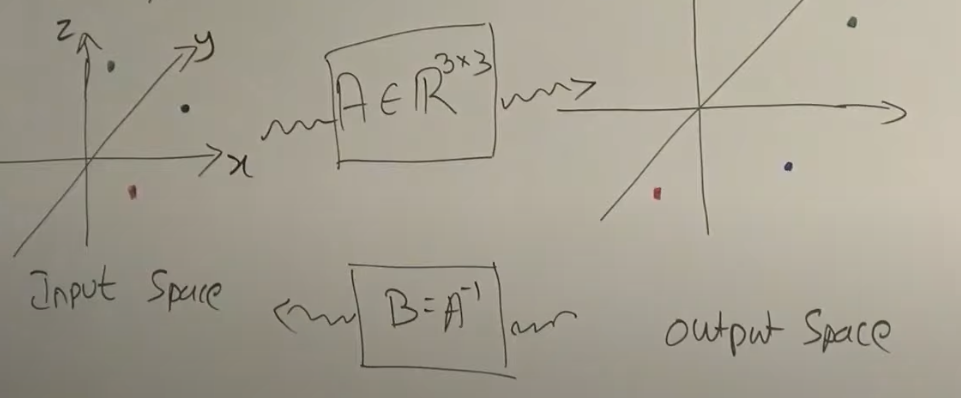
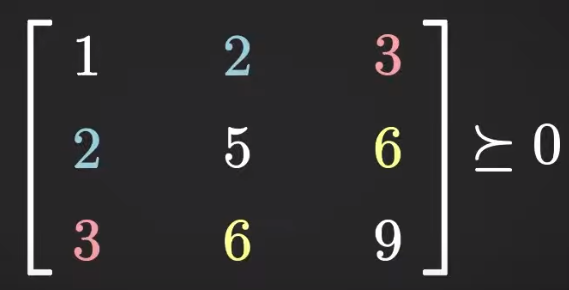
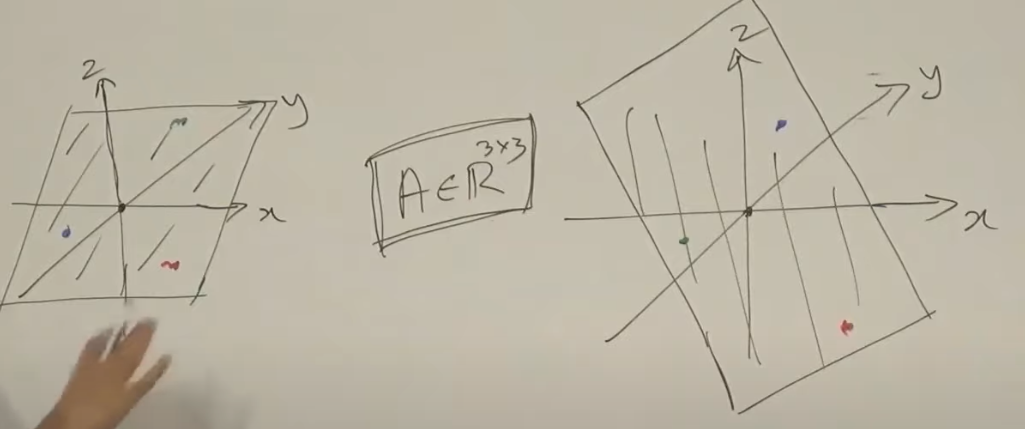
**Basics. Properties of Linear Transforms**

Full rank, unique 1:1 mapping between input and output space



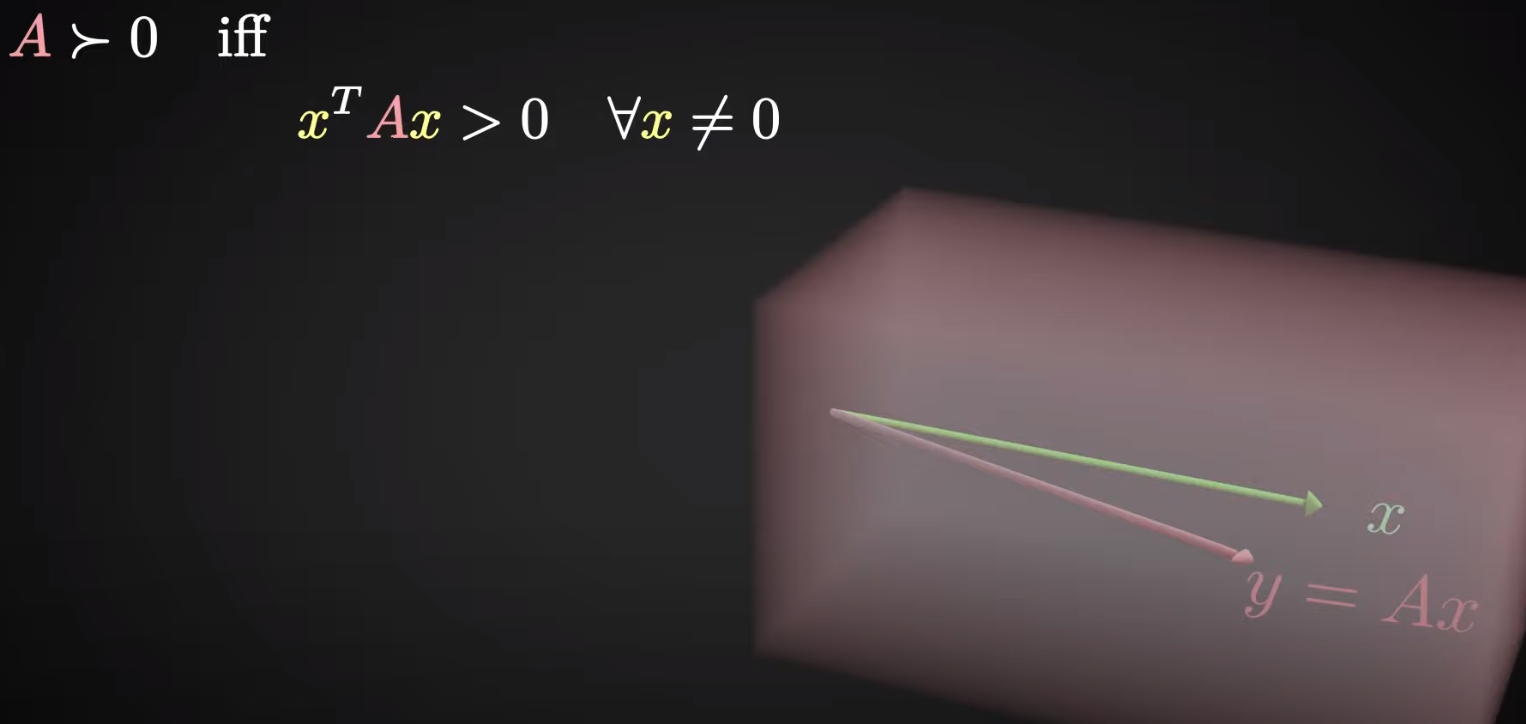
Rank deficient there exists a lower dimension subspace, which must pass through the origin, and a corresponding matrix in the output matrix that also passes through the origin, with a 1:1 mapping.

Symmetric. A = AT. Has important properties: their Eigen values are real and their eigenvectors are orthogonal 



Positive Definite. Analogous to finding if a number is greater > zero. But how can we declare if a matrix is ‘positive’?

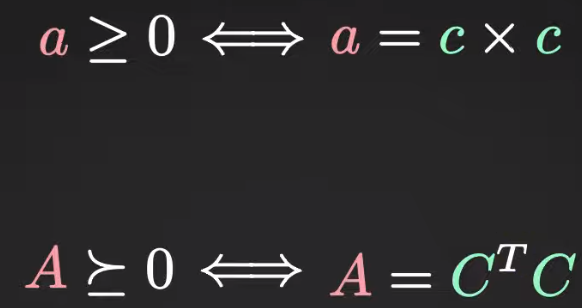
Given y=Ax, a matrix A is PD if the dot product of y and x is > 0 for all x != 0. In other words, the transformed vector y is always on the same side as x. To check if PD, need to check the eiganvalues of a matrix.



Positive Semi-Definite if the dot product can equal zero, then at least one transformed vector y exists that is orthogonal to the input vector x. Given y=Ax, a matrix A is PD if the dot product of y and x is >= 0 for all x != 0

PSD iff eigenvalues >= 0.

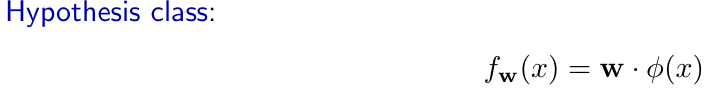
* Funky properties of PSDs, they have a ‘square root’ C. Analogous to positive scalar numbers.



PSD programming was the hottest thing in mathematical programming in the 90s. It’s a special case of convex programming.

Lyapunov function. A linear dynamical system is **asymptotically stable** iff there exists a lyapunov function!

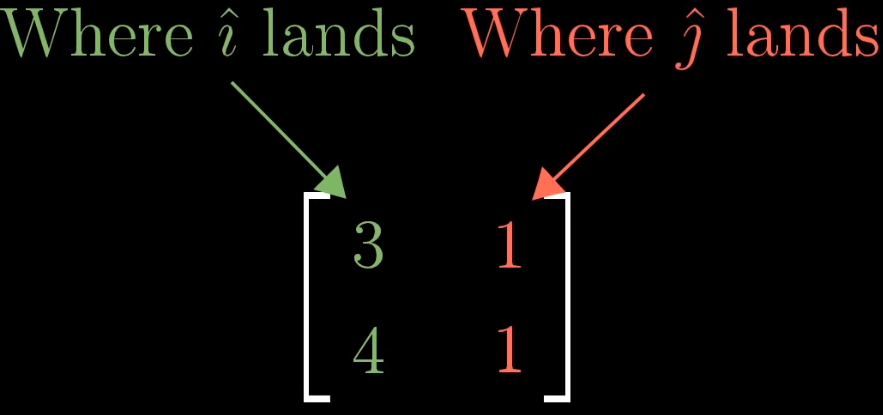
NP-Hard. No algo can solve efficiently.

**Matrix trace** Sum of all elements in the diagonal of matrix. Only defined for square nxn matrix. It is used to [measure complexity](https://www.quora.com/What-is-a-trace-as-in-trace-of-a-matrix-and-why-is-it-used) of any linear machine learning model 

Determinant: Area. properties

* determinant !=0 then there exists an inverse matrix.
* If determinant = 0.

Rank. Means the number of dimensions in the output. Also called the column space because it’s where the basis vectors land. The span of the column = the column space.



Null-space. The space of all vectors that land on zero-vector. Full rank matrix only has one solution: the origin, rank deficient matrix have lots of vectors that land on zero.

Inverse matrix: reverse a transformation to find the original vector.

**Operations**

Three types of multiplications

Vector-vector multiplication

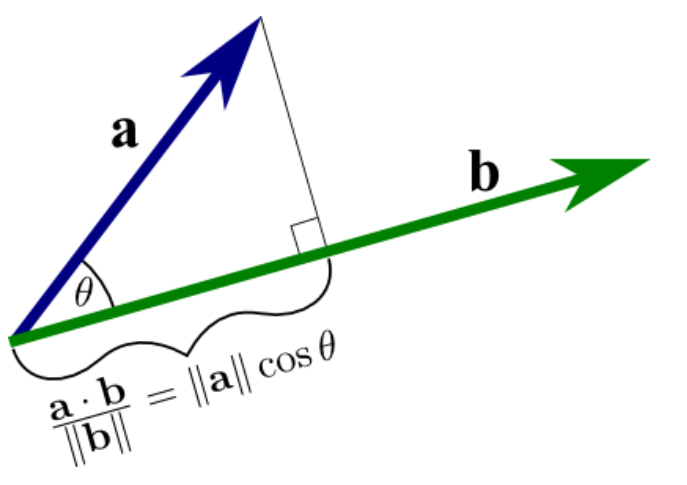
* Inner product = dot product. Result is scalar
* Outer product results in matrix. Why is it useful?

Matrix-vector multiplication

A (m x n) x(n x 1) = vector (m x1)

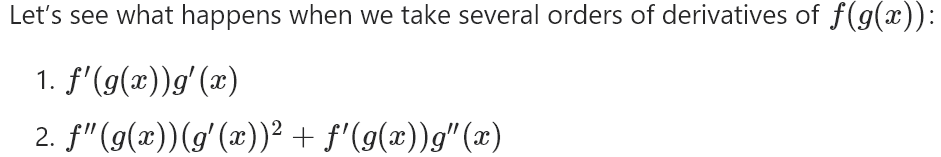
Matrix-matrix multiplication

A (m x n) x B (n x p) = C (m x p)

Projection is not the same as dot product. It’s the component of a vector.

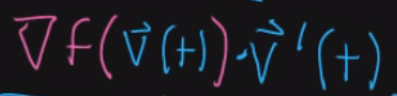
**Calculus**

Chain rule



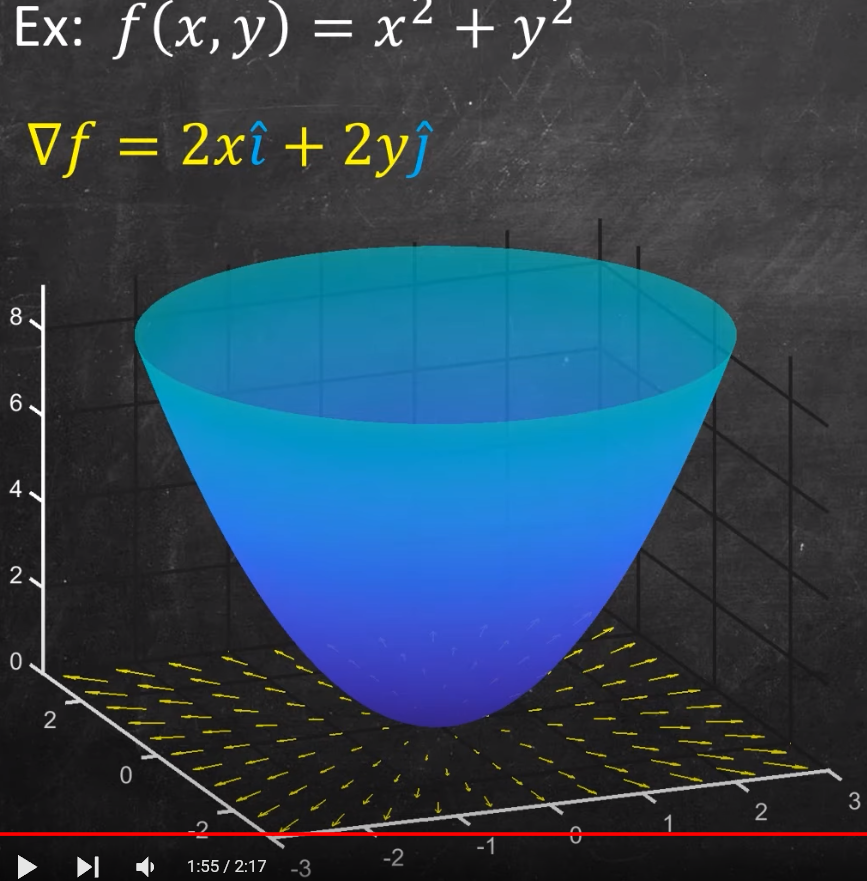
**Chain rule in vector form[[1]](#footnote-1)**

Dot product between gradient of v and vector derivative v. Interpret as the directional derivative

****

**Gradient** is a subset of a jacobian. The gradient is a vector field (a tangent vector at each pont).

****

****

When to use nabla? <https://www.youtube.com/watch?v=m2mW2FQJgEE>

Hessian

Positive Definite Hessian = Local Minimum

Negative Definite Hessian = Local Maximum

Indefinite Hessian = Saddle point

Other definiteness = Test is inconclusive

1. <https://www.youtube.com/watch?v=qZlBjnC3iro> [↑](#footnote-ref-1)