Linear algebra cheat sheet (wip) by JOL

Notations (with abridged version)

scalars $\alpha, \beta, \gamma, \delta, \varepsilon$ (real values)

vectors $\boldsymbol{b}, \boldsymbol{v}, \boldsymbol{w}, \boldsymbol{x}, \boldsymbol{y}$ $\quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = [x_1 \ \cdots \ x_n]^\top = [x_i]$

matrices A, C, M, L, U $\quad A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} = [a_{ij}]$

a *matrix* $A_{m \times n}$ with m rows and n columns is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$

Vector operations

- *usual(+,−,ext. product)* $\alpha \boldsymbol{x} + \beta \boldsymbol{y} = [\alpha x_i + \beta y_i]$
- *scalar product* $\boldsymbol{x} \cdot \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i = \boldsymbol{x}^\top \boldsymbol{y}$
- *2-norme (Euclid)* $\|\boldsymbol{x}\|_2 = \sqrt{x_1^2 + \dots + x_n^2} = \sqrt{\boldsymbol{x}^\top \boldsymbol{x}}$
- *1-norme* $\|\boldsymbol{x}\|_1 = |x_1| + \dots + |x_n|$
- ∞*-norme* $\|\boldsymbol{x}\|_\infty = \max(|x_1|, \dots, |x_n|)$
- norms mesure length of vectors
- norms are equivalent up to a $\sqrt{n}$ constant
- Euclidean norm is invariant to rigid transformation (i.e. multiplication by an orthogonal matrix).

Matrix operations

- *usual(+,−,ext. product)* $\alpha A + \beta C = [\alpha a_{ij} + \beta c_{ij}]$
- *product* if $A_{m \times n}$ and $B_{n \times p}$, then $C = AB$ is a matrix of size $m \times p$, and $c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}$
- $\begin{bmatrix} \boxed{\boldsymbol{a}_1^\top} \\ \vdots \\ \boxed{\boldsymbol{a}_m^\top} \end{bmatrix} \begin{bmatrix} \boxed{\boldsymbol{b}_1} & \cdots & \boxed{\boldsymbol{b}_p} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1^\top \boldsymbol{b}_1 & \dots & \boldsymbol{a}_1^\top \boldsymbol{b}_p \\ \vdots & \ddots & \vdots \\ \boldsymbol{a}_m^\top \boldsymbol{b}_1 & \dots & \boldsymbol{a}_m^\top \boldsymbol{b}_p \end{bmatrix} = [\boldsymbol{a}_i^\top \boldsymbol{b}_j^\top]$
- product is *associative* $A(BC) = (AB)C$,
- product is **not** *commutative* $AB \neq BA$
- *transpose* $A^\top$: $(A + B)^\top = A^\top + B^\top, (AB)^\top = B^\top A^\top$
- *conjugate transpose* A^*: same as $A^\top$ if A real-valued
- *norm* $\| \ \|_*$ *induced by vector norm* for $* \in \{1, 2, \infty\}$

$$\|A\|_* := \sup_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\|A\boldsymbol{x}\|_*}{\|\boldsymbol{x}\|_*} \text{ or (same) } \|A\|_* := \sup_{\boldsymbol{x} \neq \boldsymbol{0}, \|\boldsymbol{x}\|_* = 1} \|A\boldsymbol{x}\|_*$$

- 1-norm: *column sum* $\|A\|_1 := \max_j \sum_i |a_{ij}|$
- ∞-norm: *row sum* $\|A\|_\infty := \max_i \sum_j |a_{ij}|$
- 2-norm: *spectral row* $\|A\|_2 := \sqrt{\lambda_{\max}(A^* A)}$
- *Fröbenius norm*: $\|A\|_F := \sqrt{\operatorname{trace}(A^* A)}$
- $\|A\| \geq 0, \quad \|A\| = 0 \Leftrightarrow A = 0, \quad \|\alpha A\| = |\alpha| \ \|A\|$
- $\|A + B\| \leq \|A\| + \|B\|, \qquad \|AB\| \leq \|A\| \|B\|$

Specific matrices

- *identity matrix of size* $n \times n$ $\mathrm{Id}_n := \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}$
- $\underbrace{\begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix}}_{\text{diagonal } D} \quad \underbrace{\begin{bmatrix} a_{11} & & \\ \vdots & \ddots & \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}}_{\text{lower triangular } L} \quad \underbrace{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ & & a_{nn} \end{bmatrix}}_{\text{upper triangular } U}$
- *symmetric matrix* $A^\top = A$, *antisymmetric* $A^\top = -A$
- *orthogonal matrix* $A^\top A = AA^\top = \mathrm{Id}_n$, $\det(A) = \pm 1$, and rows/cols of A forms an orthonormal basis
- if A orthogonal and $\det(A) = 1$, then A is a rotation
- DD' is diagonal, LL' is lower triangular, UU' is upper triangular

Invertibility, determinant of matrix $A_{n \times n}$

- A is *invertible* iff (i) $\exists B, BA = \mathrm{Id}_n$, (ii) $\det(A) \neq 0$, (iii) $\operatorname{rank}(A) = n$, or (iv) $A\boldsymbol{x} = \boldsymbol{0}$ has only $\boldsymbol{x} = \boldsymbol{0}$ as solution.
- $\det([a_{11}]) = a_{11}$, $\det(A) = \sum_i a_{ij} \Delta_{ij}$, where *cofactor* $\Delta_{ij} := (-1)^{i+j} \det(A_{ij})$, if A_{ij} is the matrix A without row i and column j.
- if A is invertible, $A^{-1} = \frac{[\Delta_{ij}]}{\det(A)}$, $\Rightarrow$ solve $A\boldsymbol{x} = \boldsymbol{b}$

Cramer's rule $x_i = \det\left(\begin{bmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{bmatrix} \right) / \det(A)$

- complexity of above is $\mathcal{O}(n!)$, used when $n \leq 4$
- determinant of diagonal or triangular matrix is product of diagonal elements
- matrix condition $K_p(A) = \|A\|_p \|A^{-1}\|_p$, $K_2(A) = \frac{|\sigma_{\max}(A)|}{|\sigma_{\min}(A)|}$
- $K(A) \geq K(\mathrm{Id}_n) = 1$ for any p-norm

Solving linear systems $A\boldsymbol{x} = \boldsymbol{b}$

- if A is "L", easy to solve $L\boldsymbol{x} = \boldsymbol{b}$ in n^2 flops
- if A is "U", easy to solve $U\boldsymbol{x} = \boldsymbol{b}$ in n^2 flops
- "LU" factorization: find L, U with $A = LU$,
  - L and U found by Gauss elimination in $\frac{2}{3}n^3$ flops
  - solve $L\boldsymbol{y} = \boldsymbol{b}$ then $U\boldsymbol{x} = \boldsymbol{y}$
- "PLU" factorization: add permutation mat. P
  - $PA = LU$ (with $P^\top = P^{-1}$)
  - look for the biggest pivot
  - A invertible $\Leftrightarrow$ A has a PLU factorization
- Cholesky factorization
  - A **SPD**: $A = A^\top$ and $\forall \boldsymbol{x} \neq 0, \boldsymbol{x}^\top A \boldsymbol{x} > 0$
  - find L with $A = LL^\top$ in $\frac{1}{3}n^3$ flops
  - faster and more stable than PLU
  - variant as $A = LDL^\top$

Numerical stability of $\tilde{f}$ wrt f

- machine precision $\mathfrak{u}$ ($\approx 1e - 16$ with double)
- forward error $\|\tilde{f}(d) - f(d)\| / \|f(d)\|$ ("aval")
- backward error $\|\tilde{d} - d\| / \|d\|$, where $f(\tilde{d}) = \tilde{f}(d)$
- *forward stability*: forward error is $\mathcal{O}(u)$, e.g. fl, $\otimes$
- *backward stability*: backward error is $\mathcal{O}(u)$, e.g. std operations, scalar/matrix product, etc
- forward error = $K_{\mathrm{rel}}(f) \times$ backward error
- forward stable $\Rightarrow$ backward stable