<div align="center">

Mathematics Department Stanford University
Summary of Math 51H Linear Algebra Material

</div>

The following is a brief summary of the main results covered in the linear algebra part of 51H; you should of course know all these results and their proofs and be able to apply them in the manner required e.g. as in the homework problems.

Vectors in $\mathbb{R}^n$, meaning of linear combinations, l.i., l.d., span, subspace. Dot product, Cauchy Schwarz inequality and its proof. Angle between non-zero vectors.

Gaussian elimination, Underdetermined systems lemma. The Linear Dependence Lemma and its consequences: Basis Theorem and definition of dimension, and the facts that (a) k l.i. vectors in a k-dimensional subspace automatically span, (b) k vectors which span a k-dimensional subspace are automatically l.i.

For an $m \times n$ matrix A: Definition of $N(A)$, $C(A)$. Rank/nullity theorem, Basic matrix algebra (product and sums) and the fact that $\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}$. The transpose A^{T} of A and the formula $(AB)^{\mathrm{T}} = B^{\mathrm{T}} A^{\mathrm{T}}$. Reduction of A to reduced row echelon form rref A and its consequences, including the alternate proof of the rank nullity theorem and the fact that $C(A)$ is the span of the columns of A with column numbers equal to the column numbers of the pivot columns of rref A.

For each subspace V of $\mathbb{R}^n$, the definition of the orthogonal complement $V^{\perp}$ of V, and the facts that $V \cap V^{\perp} = \{\underline{0}\}$, $V + V^{\perp} = \mathbb{R}^n$ and that this is a direct sum (i.e. for each $\underline{z} \in \mathbb{R}^n$ there are *unique* $\underline{x}, \underline{y} \in V, V^{\perp}$ with $\underline{z} = \underline{x} + \underline{y}$), $\dim V + \dim V^{\perp} = n$, $(V^{\perp})^{\perp} = V$. Existence of unique orthogonal projection $P : \mathbb{R}^n \to \mathbb{R}^n$ with the properties (a) $P(\underline{x}) \in V$ and (b) $\underline{x} - P(\underline{x}) \in V^{\perp} \forall \underline{x} \in \mathbb{R}^n$. Proof that such P is automatically has the additional properties: (i) It is linear, (ii) $P(\underline{x}) = \underline{x} \forall \underline{x} \in V$, (iii) it is symmetric (i.e. $\underline{x} \cdot P(\underline{y}) = \underline{y} \cdot P(\underline{x}) \forall \underline{x}, \underline{y} \in \mathbb{R}^n$), (iv) $P(V^{\perp}) = \{\underline{0}\}$ and (v) $\|\underline{x} - P(\underline{x})\|$ gives the distance of a point $\underline{x}$ from V (i.e. $P(\underline{x})$ is the nearest point of the subspace V to the vector $\underline{x}$). (Terminology: P is called "the orthogonal projection onto V.")

The fact that $C(A^{\mathrm{T}}) = (N(A))^{\perp}$ for any $m \times n$ matrix, and the consequence that row $\text{rank}(A) = \text{rank}(A) = \text{rank}(A^{\mathrm{T}})$.

Affine spaces $\underline{x}_0 + V$ in $\mathbb{R}^n$ (where V is a subspace) and the fact that the nearest point of $\underline{x}_0 + V$ to $\underline{0}$ is given by $\underline{x}_0 - P(\underline{x}_0)$, where P is the orthogonal projection onto V.

The main theorem of inhomogeneous systems: that if A is $m \times n$ and $\underline{y} \in \mathbb{R}^m$ is given, then
   (i)   If $A\underline{x} = \underline{y}$ has at least one solution $\underline{x}_0$, then the whole solution set is precisely the affine space $\underline{x}_0 + N(A)$,
   (ii) $A\underline{x} = \underline{y}$ has a solution $\iff \underline{y} \in C(A) \iff \underline{y} \in (N(A^{\mathrm{T}}))^{\perp}$

Permutations and definition of even/odd permutations; inverse permutation of a given permutation and the fact that the parity of a permutation and its inverse are the same.

Definition of determinant of an $n \times n$ matrix (in terms of the function D and the formula for $\det A$ as a sum of $n!$ terms, each $\pm$ a product of n terms, each term taken from a distinct row and column of A); the properties that (a) $\det(A)$ is linear in each row, (b) $\det(\tilde{A}) = -\det(A)$ if $\tilde{A}$ is obtained by interchanging two distinct rows of A, and (c) $\det(A) = \det(A^{\mathrm{T}})$. Computation of $\det A$ by elementary row operations, and the fact that $\det A \neq 0 \iff \text{rref} A = I \iff \text{rref} A$ has no zero rows. The formulae for the expansion of $\det A$ along the i'th row and j'th column of A and the corresponding formulae $\sum_{j=1}^{n}(-1)^{i+j}a_{kj}\det(A_{ij}) = \det A \, \delta_{ik}$ for each $i, k = 1, \ldots, n$, $\sum_{i=1}^{n}(-1)^{i+j}a_{ik}\det(A_{ij}) = \det A \, \delta_{jk}$ for each $j, k = 1, \ldots, n$, where δ_{ij} is the i, j'th entry of the identity matrix (i.e. 1 if $i = j$ and 0 if $i \neq j$). The formula $A^{-1} = (\det A)^{-1}\big((-1)^{i+j}\det(A_{ji})\big)$ if $\det(A) \neq 0$. Computation of A^{-1} via elementary row operations.

For an $n \times n$ matrix A: A^{-1} exists $\iff \det A \neq 0 \iff N(A) = \{\underline{0}\} \iff \text{rank}(A) = n \iff \text{rref}(A) = I \iff$ the map $\underline{x} \mapsto A\underline{x}$ is 1:1 $\iff$ the map $\underline{x} \mapsto A\underline{x}$ is onto. The formula $(AB)^{-1} = B^{-1}A^{-1}$ if A^{-1}, B^{-1} exist and if B is $n \times n$.

Gram-Schmidt orthogonalization and the existence of an orthonormal basis for each non-trivial subspace V of $\mathbb{R}^n$; the explicit formula for the orthogonal projection P onto V: $P(\underline{x}) = \sum_{j=1}^{k}(\underline{x} \cdot \underline{w}_j)\underline{w}_j$, where $\underline{w}_1, \ldots, \underline{w}_k$ is any orthonormal basis for the non-trivial k-dimensional subspace V, and the formula matrix of $P = WW^{\mathrm{T}}$, where W is the $n \times k$ matrix with j'th column $= \underline{w}_j$.

Definition of eigenvalues/eigenvectors of an $n \times n$ matrix.

The Spectral Theorem (that if A is a symmetric matrix, then there is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A, and if Q is the matrix with columns given by such an orthonormal basis, then Q is orthogonal (i.e. $Q^{\mathrm{T}}Q = I$) and $Q^{\mathrm{T}}AQ = $ diagonal matrix with the eigenvalues of A along the leading diagonal).

The following is a brief summary of the main results covered in the multivariable calculus and real analysis part of 51H; you should of course know all these results and their proofs and be able to apply them in the manner required e.g. in the homework assignments.

Open and closed sets in $\mathbb{R}^n$. Theorem that a set C is closed if and only if its complement $\mathbb{R}^n \setminus C$ is open. (Equivalently, since $\mathbb{R}^n \setminus (\mathbb{R}^n \setminus C) = C$, a set U is open if anly only if its complement $\mathbb{R}^n \setminus U$ is closed). Bolzano-Weierstrass theorem for bounded sequences in $\mathbb{R}^n$. Theorem that a continuous real-valued function on a compact set attains both its maximum and minimum values.

Definition of differentiability, and the fact that differentiability of f implies all partials and all directional derivatives exist, and $D_{\underline{v}} f(\underline{a}) = \sum_{j=1}^{n} v_j D_j f(\underline{a})$ if f is differentiable at $\underline{a}$.

The chain rule for the composite of differentiable functions. Theorem that differentiability at point $\underline{a}$ implies continuity at $\underline{a}$.

Theorem that f of class C^1 on U implies f differentiable at each point of U, and f of class C^2 on U implies $D_i D_j f = D_j D_i f$ at each point of U. The gradient ∇f of a real-valued C^1 function f and the fact that the gradient gives the direction of fastest increase of f at points where $\nabla f \neq \underline{0}$.

Quadratic forms $Q(\xi)$ on $\mathbb{R}^n$ and definition of positive definite and negative definite. The fact that Q positive definite implies that there is an $m > 0$ such that $Q(\xi) \geq m\|\xi\|^2$ for all $\xi \in \mathbb{R}^n$.

For a C^2 function on the ball $B_\rho(\underline{x}_0)$, the second derivative identity $f(\underline{x}) = f(\underline{x}_0) + (\underline{x} - \underline{x}_0) \cdot \nabla f(\underline{x}_0) + \frac{1}{2} Q_{\underline{x}_0}(\underline{x} - \underline{x}_0) + E(\underline{x})$, with $\lim_{\underline{x} \to \underline{x}_0} \|\underline{x} - \underline{x}_0\|^{-2} E(\underline{x}) = 0$, where $Q_{\underline{x}_0}(\xi)$ is the Hessian quadratic form $\sum_{i,j=1}^{n} D_i D_j f(\underline{x}_0) \xi_i \xi_j$ of f at $\underline{x}_0$. The consequent facts that if $\underline{x}_0$ is a critical point (i.e. $\nabla f(\underline{x}_0) = \underline{0}$) then (i) if the Hessian quadratic form $Q_{\underline{x}_0}(\xi)$ is positive definite, then f has a local minimum at $\underline{x}_0$, and (ii) if $Q_{\underline{x}_0}(\xi)$ is negative definite, then f has a local maximum $\underline{x}_0$.

Length of a C^0 curve $\gamma : [a,b] \to \mathbb{R}^n$, and the fact the C^1 curves have finite length given by the formula $\ell(\gamma) = \int_a^b \|\gamma'(t)\| \, dt$. Arc-length parameter $s = S(t)$ for C^1 curves $\gamma(t)$ with $\gamma'(t) \neq 0$, velocity and unit tangent vectors for such curves, and the curvature vector of such a curve assuming γ is C^2.

Definition of k-dimensional C^1 submanifold M of $\mathbb{R}^n$ and the tangent space $T_{\underline{a}} M$. Proof that the tangent space can be expressed as the span of the partial derivatives of the relevant graph map. Tangential gradient $\nabla_M f$ of a C^1 function f defined in a neighborhood of M and the fact that $\nabla_M f = \underline{0}$ at a local max/min of $f|M$. Lagrange muliplier theorem as in 9.14 of Ch.2 of text.

Contraction mapping principle and its proof. The inverse function theorem and its proof. Implicit function theorem. (Note: Implicit function theorem will not be tested in the final examination)

Real Analysis Lecture 1: Basic properties of the real numbers, including the supremum axiom and the density of the rationals and the irrationals.

Real Analysis Lecture 2: Basic properties of real sequences, including definition of convergence and the Bolzano-Weierstrass theorem for bounded sequences in $\mathbb{R}$, and the fact that bounded monotone sequences are convergent.

Real Analysis Lecture 3: Definition and basic properties of continuous functions, including the proof that a continuous function on a closed interval attains its maximum and minimum values.

Real Analysis Lecture 4: Basic properties of series of real numbers, including definition of convergence, the theorem that convergence implies n'th term $\to 0$. For sequences with non-negative terms, the theorem that the series converges if and only if the sequence of partial sums is bounded. Comparison and integral tests for convergence/divergence. Absolute convergence implies convergence. (Omit rearrangement of series, since we did not cover that topic this year.)

Real Analysis Lecture 5: Power series. Definition of radius of convergence and theorem that for a given power series $\sum_{n=0}^{\infty} a_n x^n$ there are just 3 possibilities: (i) the series diverges at each point $x \neq 0$, (ii) the series converges absolutely at each point of $\mathbb{R}$, or (iii) there is $\rho > 0$ such that the series converges absolutely for each point x with $|x| < \rho$, and diverges for each point x with $|x| > \rho$.

Real Analysis Lecture 6: Taylor series: Change of base point theorem, termwise differentiability theorem, Taylor's theorem, and the sufficient condition $\sup_{n=0,1,2,\ldots} \sup_{|x|<r} \frac{|f^{(n)}(x)|}{n!} r^n \leq M$ to ensure convergence of the Taylor series to the function f on the interval $|x| < r$.