Mathematics Department Stanford University
Math 51H Final Examination, December 8, 2014

2 Hours

Unless otherwise indicated, you can use results covered in lecture and homework, provided they are clearly stated.

If necessary, continue solutions on backs of pages
Note: work sheets are provided for your convenience, but will not be graded

Question 6b is extra credit only! Work on it only if you are done with the other problems!

Q.1	
Q.2	
Q.3	
Q.4	
Q.5	
Q.6a	
T/35	
Q.6b	

Name (Print Clearly): ____________________

I understand and accept the provisions of the honor code (Signed) ____________________

1(a) (3 points): Find an orthonormal basis for the subspace of $\mathbb{R}^4$ spanned by the vectors $(1,0,2,0)^{\mathrm{T}}, (1,0,0,3)^{\mathrm{T}}, (0,2,0,1)^{\mathrm{T}}$, and write down an explicit formula (involving numbers and matrix operations only) for the matrix of the orthogonal projection to this subspace (but you do not need to compute it).

Solution: We apply the Gram-Schmidt algorithm. Since the first and last vectors are already orthogonal, reorder them as $v_1 = (1,0,2,0)^{\mathrm{T}}, v_2 = (0,2,0,1)^{\mathrm{T}}, v_3 = (1,0,0,3)^{\mathrm{T}}$ to simplify the calculation. Then v_1 and v_2 are orthogonal, so we only need to make them unit length, i.e. replace them by $w_1 = \frac{1}{\sqrt{5}}v_1$ and $w_2 = \frac{1}{\sqrt{5}}v_2$. Finally, we need to replace v_3 by its projection to the orthocomplement of $\mathrm{Span}\{w_1, w_2\}$, and make the result unit length by dividing by its length. The first step gives

$$v_3-(w_1\cdot v_3)w_1-(w_2\cdot v_3)w_2 = (1,0,0,3)^{\mathrm{T}} - \frac{1}{5}(1,0,2,0)^{\mathrm{T}} - \frac{3}{5}(0,2,0,1)^{\mathrm{T}} = (\frac{4}{5}, -\frac{6}{5}, -\frac{2}{5}, \frac{12}{5})^{\mathrm{T}};$$

the length of this vector is $\frac{1}{5}\sqrt{16+36+4+144} = \frac{1}{5}\sqrt{200} = 2\sqrt{2}$, so we get

$$w_1 = (\frac{1}{\sqrt{5}}, 0, \frac{2}{\sqrt{5}}, 0)^{\mathrm{T}},\ w_2 = (0, \frac{2}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}})^{\mathrm{T}},\ w_3 = (\frac{2}{5\sqrt{2}}, -\frac{3}{5\sqrt{2}}, -\frac{1}{5\sqrt{2}}, \frac{6}{5\sqrt{2}})^{\mathrm{T}}.$$

The orthogonal projection to the span of w_1, w_2, w_3 of a vector v is

$$Pv = (v\cdot w_1)w_1 + (v\cdot w_2)w_2 + (v\cdot w_3)w_3 = \sum_{j=1}^{3} w_j(w_j^{\mathrm{T}}v) = (\sum_{j=1}^{3} w_j w_j^T)v$$

so the matrix is

$$\begin{aligned}\sum_{j=1}^{3} w_j w_j^T =& (\frac{1}{\sqrt{5}}, 0, \frac{2}{\sqrt{5}}, 0)^{\mathrm{T}}(\frac{1}{\sqrt{5}}, 0, \frac{2}{\sqrt{5}}, 0) + (0, \frac{2}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}})^{\mathrm{T}}(0, \frac{2}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}}) \\ &+ (\frac{2}{5\sqrt{2}}, -\frac{3}{5\sqrt{2}}, -\frac{1}{5\sqrt{2}}, \frac{6}{5\sqrt{2}})^{\mathrm{T}}(\frac{2}{5\sqrt{2}}, -\frac{3}{5\sqrt{2}}, -\frac{1}{5\sqrt{2}}, \frac{6}{5\sqrt{2}}).\end{aligned}$$

(b) (2 points) Suppose A is an $m \times n$ matrix. Prove that the dimensions of $C(A)$ and $C(A^{\mathrm{T}})$ are the same.

Solution: Since $N(A)^{\perp} = C(A^{\mathrm{T}})$, with both being subspaces of $\mathbb{R}^n$, we have $\dim C(A^{\mathrm{T}}) = n - \dim N(A)$. On the other hand, by the rank nullity theorem $\dim N(A) + \dim C(A) = n$ since A represents a linear operator $\mathbb{R}^n \to \mathbb{R}^m$. Thus, $n - \dim N(A) = \dim C(A)$, so $\dim C(A) = \dim C(A^{\mathrm{T}})$.

2(a) (4 points): For $(x,y,z) \in \mathbb{R}^3$, let $f(x,y,z) = \frac{16}{3}x^2 + z$ and let $S = \{(x,y,z) \in \mathbb{R}^3 : x^2 + y^4 + z^6 = 1\}$. (i) Show that $f|_S$ attains a maximum and a minimum value. (ii) Find a point where each of these is attained.

Solution: (i) If $(x,y,z) \in S$, then $x^2 + y^4 + z^6 = 1$ shows that $x^2, y^4, z^6 \leq 1$ since all summands are non-negative. Thus, $|x|, |y|, |z| \leq 1$, so S is bounded. On the other hand, the map $g(x,y,z) = x^2 + y^4 + z^6 - 1$ defined on $\mathbb{R}^3$ is C^∞, so in particular continuous, so $g^{-1}(\{0\}) = S$ is closed since $\{0\} \subset \mathbb{R}$ is closed. Correspondingly S is compact (as it is closed and bounded), so any continuous function, such as $f|_S$, attains its maximum and minimum on S. (ii) We have g is C^∞ and $Dg(x,y,z) = (2x, 4y^3, 6z^5)$, so the vanishing of Dg means $x = y = z = 0$, and thus Dg does not vanish on S. Correspondingly, by the implicit function theorem, S is a C^∞ submanifold of $\mathbb{R}^3$. Further, any critical points p of $f|_S$, which includes all local maxima and minima, satisfy that $Df(p) = \lambda Dg(p)$ for some $\lambda \in \mathbb{R}$ by the Lagrange multiplier theorem. Since $Df(x,y,z) = (\frac{32}{3}x, 0, 1)$, this means that $\frac{32}{3}x = 2\lambda x$, $0 = 4\lambda y^3$, $1 = 6\lambda z^5$. The last of these shows $\lambda \neq 0$, so from the second of these $y = 0$, and from the first either $x = 0$ or $\lambda = \frac{16}{3}$. If $x = 0$, we get (as $y = 0$), $z = \pm 1$; if $x \neq 0$ then $1 = 32z^5$, so $z = \frac{1}{2}$, and thus $x = \pm\sqrt{1 - \frac{1}{64}} = \pm\frac{\sqrt{63}}{8}$. So the critical points are $(0,0,1)$, $(0,0,-1)$, $(\frac{\sqrt{63}}{8}, 0, \frac{1}{2})$ and $(-\frac{\sqrt{63}}{8}, 0, \frac{1}{2})$. The maxima and minima must be among these points; and f takes the values $f(0,0,1) = 1$, $f(0,0,-1) = -1$, $f(\frac{\sqrt{63}}{8}, 0, \frac{1}{2}) = f(-\frac{\sqrt{63}}{8}, 0, \frac{1}{2}) = \frac{16 \cdot 63}{3 \cdot 64} + \frac{1}{2} = \frac{21}{4} + \frac{1}{2} > 1$. Thus, the maximum is attained at $(\pm\frac{\sqrt{63}}{8}, 0, \frac{1}{2})$ and the minimum at $(0,0,-1)$.

(b) (3 points): Find the determinant and the inverse of the matrix

$$A = \begin{pmatrix} 3 & 6 & -9 \\ 0 & 2 & 4 \\ 0 & 0 & 1 \end{pmatrix}.$$

Solution: Since the matrix is upper triangular, its determinant is the product of its diagonal entries, i.e. 6. As for the inverse,

$$\left(\begin{array}{ccc|ccc} 3 & 6 & -9 & 1 & 0 & 0 \\ 0 & 2 & 4 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \begin{array}{c} r_1 \mapsto \frac{1}{3}r_1 \\ r_2 \mapsto \frac{1}{2}r_2 \end{array} \left(\begin{array}{ccc|ccc} 1 & 2 & -3 & \frac{1}{3} & 0 & 0 \\ 0 & 1 & 2 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \begin{array}{c} r_1 \mapsto r_1 + 3r_3 \\ r_2 \mapsto r_2 - 2r_3 \end{array} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & \frac{1}{3} & 0 & 3 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & -2 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$r_1 \mapsto r_1 - 2r_2 \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{3} & -1 & 7 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & -2 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

so the inverse is $\begin{pmatrix} \frac{1}{3} & -1 & 7 \\ 0 & \frac{1}{2} & -2 \\ 0 & 0 & 1 \end{pmatrix}$.

3(a) (3 points): Suppose $U \subset \mathbb{R}^n$ is open, $\underline{x}_0 \in U$. (i) State the definition of a map $f : U \to \mathbb{R}^m$ being differentiable at $\underline{x}_0$. (ii) Show that if $f : U \to \mathbb{R}^m$ is differentiable at $\underline{x}_0$ then it is continuous at $\underline{x}_0$.

Solution: f is differentiable at $\underline{x}_0$ with derivative $A = (Df)(\underline{x}_0)$, a linear map $A : \mathbb{R}^n \to \mathbb{R}^m$, if for all $\varepsilon > 0$ there is $\delta > 0$ such that $B_\delta(\underline{x}_0) \subset U$ and such that $\|\underline{x} - \underline{x}_0\| < \delta$ implies $\|f(\underline{x}) - f(\underline{x}_0) - A(\underline{x} - \underline{x}_0)\| \leq \varepsilon\|\underline{x} - \underline{x}_0\|$.

If f is differentiable at $\underline{x}_0$ with derivative A, then given any $\tilde{\varepsilon} > 0$ there is $\tilde{\delta} > 0$ such that $B_{\tilde{\delta}}(\underline{x}_0) \subset U$ and such that $\|\underline{x} - \underline{x}_0\| < \tilde{\delta}$ implies $\|f(\underline{x}) - f(\underline{x}_0) - A(\underline{x} - \underline{x}_0)\| \leq \tilde{\varepsilon}\|\underline{x} - \underline{x}_0\|$. Thus, for $|\underline{x} - \underline{x}_0\| < \tilde{\delta}$,

$$\|f\underline{x}) - f(\underline{x}_0)\| \leq \|f(\underline{x}) - f(\underline{x}_0) - A(\underline{x} - \underline{x}_0)\| + \|A(\underline{x} - \underline{x}_0)\| \leq \tilde{\varepsilon}\|\underline{x} - \underline{x}_0\| + \|A\|\|x - \underline{x}_0\| = (\|A\| + \tilde{\varepsilon})\|x - \underline{x}_0\|.$$

Now, use this with $\tilde{\varepsilon} = 1$ to conclude that $|\underline{x} - \underline{x}_0\| < \tilde{\delta}$ implies $\|f(\underline{x}) - f(\underline{x}_0)\| \leq (\|A\| + 1)\|\underline{x} - \underline{x}_0\|$. For any specified $\varepsilon > 0$, the right hand side will be $< \varepsilon$ if $\|\underline{x} - \underline{x}_0\| < \frac{\varepsilon}{\|A\|+1}$; since we needed in addition that $\|\underline{x} - \underline{x}_0\| < \tilde{\delta}$ to get this inequality in the first place, we conclude that f is coninuous at $\underline{x}_0$ by taking, for given $\varepsilon > 0$, $\delta = \min(\tilde{\delta}, \frac{\varepsilon}{\|A\|+1})$.

3(b) (4 points): Show the intermediate value theorem: if $f : [a, b] \to \mathbb{R}$ is continuous, $f(a) = \alpha$, $f(b) = \beta$, $\alpha < c < \beta$ then there exists $x \in (a, b)$ such that $f(x) = c$.
Hint: Consider $\inf\{z \in [a, b] : \ f(z) > c\}$.

Solution: As suggested by the hint, let $x = \inf S$, $S = \{z \in [a, b] : \ f(z) > c\}$. Note that S is non-empty, as b is in it, and it is bounded below since a is a lower bound, so this inf exists. Further, $x > a$ since by the continuity of f at a and as $c > \alpha$ so $\varepsilon = c - \alpha > 0$ there exists $\delta > 0$ such that $z < a + \delta$, $z \in [a, b]$, implies $f(z) < f(a) + \varepsilon = c$, i.e. $[a, a + \delta) \cap S = \emptyset$, and thus $x \geq a + \delta$ (since any element of $[a, a + \delta)$ is a lower bound for S). Since f is continuous, it is continuous at x in particular, so if x_n is any sequence in $[a, b]$ with $\lim x_n = x$ then $\lim f(x_n) = f(x)$ as continuity implies sequential continuity. Since $x > a$, we can take a sequence $x_n \in [a, b]$ with $x_n < x$ for all n and such that $x_n \to x$, e.g. take $x_n = \max(a, x - 1/n)$. Then $x_n \notin S$ (since x is a lower bound for S), so $f(x_n) \leq c$, so we conclude that $f(x) = \lim f(x_n) \leq c$ as well. On the other hand, since $x = \inf S$, there exists a sequence $\{x_n\}_{n=1}^{\infty}$ in S converging to x, so $f(x_n) > c$, and so $f(x) = \lim f(x_n) \geq c$. Combining these two, we deduce that $f(x) = c$, completing the proof.

4(a) (3 points): Let $A = (a_{ij})$ be an $n \times n$ symmetric matrix. Prove that the quadratic form $\mathcal{A}(\underline{x}) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$ is positive definite $\iff$ all the eigenvalues of A are positive.
Hint: Spectral Theorem.

Solution: By the spectral theorem there is an orthonormal basis $\underline{v}_1, \ldots, \underline{v}_n$ of $\mathbb{R}^n$ consisting of eigenvectors of A with eigenvalue λ_j: $A\underline{v}_j = \lambda_j \underline{v}_j$. Now, the quadratic form is $\mathcal{A}(\underline{x}) = \underline{x} \cdot A\underline{x}$. Writing $\underline{x} = \sum_{j=1}^{n} c_j \underline{v}_j$ (which can be done by the basis property), we have

$$\mathcal{A}(\underline{x}) = (\sum_{j=1}^{n} c_j \underline{v}_j) \cdot A(\sum_{k=1}^{n} c_k \underline{v}_k) = \sum_{j,k=1}^{n} c_j c_k v_j \cdot A v_k = \sum_{j,k=1}^{n} c_j c_k \lambda_k v_j \cdot v_k = \sum_{j=1}^{n} \lambda_j c_j^2.$$

Now, $\mathcal{A}$ is positive definite means exactly that $\mathcal{A}(\underline{x}) > 0$ if $\underline{x} \neq 0$, which in turn is equivalent to $\sum_{j=1}^{n} \lambda_j c_j^2 > 0$ if $(c_1, \ldots, c_n) \neq 0$, i.e. not all c_j vanish. This immediately gives that if all eigenvalues λ_j of A are positive then $\mathcal{A}$ is positive definite (since $\lambda_j c_j^2 \geq 0$, and at least for one j it is > 0, so the sum is > 0). Conversely, if A is not positive definite, then there is at least one k such that $\lambda_k \leq 0$. Taking $c_k = 1$, $c_j = 0$ if $j \neq k$, we have $\sum_{j=1}^{n} \lambda_j c_j^2 = \lambda_k \leq 0$, so $\mathcal{A}$ is not positive definite, completing the proof.

(b) (3 points): Find all critical points of the map $f : \mathbb{R}^3 \to \mathbb{R}$ defined by $f(x, y, z) = \frac{1}{5}(x^5 + y^5) + \frac{1}{3}z^3 - x - y - 4z$, and discuss whether these points are local maxima/minima for f. Justify all claims with proofs, possibly using theorems from lecture.

Solution: Since $Df(x, y, z) = (x^4 - 1, y^4 - 1, z^2 - 4)$, the critical points of f satisfy $x^4 - 1 = 0$, $y^4 - 1 = 0$, $z^2 = 4$; this gives $x = \pm 1$, $y = \pm 1$, $z = \pm 2$. Further, the Hessian is

$$\begin{pmatrix} 4x^3 & 0 & 0 \\ 0 & 4y^3 & 0 \\ 0 & 0 & 2z \end{pmatrix},$$

so it has eigenvalues given by the diagonal entries, which are ± 4, ± 4 and ± 4 at these critical points. In order to have a local maximum one must have eigenvalues ≤ 0, and one definitely does have a local maximum if the eigenvalues are < 0; analogously with local minima. Thus, the only possible local maximum is $(-1, -1, -2)$, and this is indeed a local maximum, while the only possible local minimum is $(1, 1, 2)$, which indeed is a local minimum.

5(a) (3 points): Suppose that $\{M_n\}_{n=1}^\infty$ is a sequence with $M_n \geq 0$ for all n, and $\sum_{n=1}^\infty M_n$ converges. Show that if $\{x_n\}_{n=1}^\infty$ is a sequence in a complete normed vector space $(V, \|.\|)$ (if you wish, you may take $V = \mathbb{R}^m$ with the standard norm) and $\|x_n\| \leq M_n$ for all n, then $\sum_{n=1}^\infty x_n$ converges in V, i.e. $\lim_{k\to\infty} \sum_{n=1}^k x_n$ exists. (This is the Weierstrass M-test.)

Solution: Since $\|x_n\| \leq M_n$ for all n, $\sum_{n=1}^\infty \|x_n\|$ converges (in $\mathbb{R}$) since it is a series with non-negative terms, and its partial sums are bounded by those of $\sum_{n=1}^\infty M_n$, which are in turn bounded by their limit, $\sum_{n=1}^\infty M_n$.

Now, we claim that the partial sums $s_n = \sum_{k=1}^n x_k$ form a Cauchy sequence in V. If we show this, the completeness of V implies that they converge, i.e. that $\sum_{k=1}^\infty x_k$ converges, completing the proof.

But if $n > m$ (with $n < m$ following by relabelling, and $n = m$ being automatic),

$$\|s_n - s_m\| = \|\sum_{k=m+1}^n x_k\| \leq \sum_{k=m+1}^n \|x_k\|.$$

On the other hand, if $\sigma_n = \sum_{k=1}^n \|x_k\|$, then $\sigma_n - \sigma_m = \sum_{k=m+1}^n \|x_k\|$ as well. Since $\{\sigma_n\}_{n=1}^\infty$ converges, it is Cauchy, so given $\varepsilon > 0$ there is N such that $n, m \geq N$ implies $|\sigma_n - \sigma_m| < \varepsilon$. Thus, for $n, m \geq N$, $\|s_n - s_m\| \leq |\sigma_n - \sigma_m| < \varepsilon$, proving the claimed Cauchy property, and thus completing the proof.

(b) (4 points): Suppose that $A : \mathbb{R}^m \to \mathbb{R}^m$ linear satisfies $\|A\| < 1$. Show that $(I-A)^{-1} = I + \sum_{n=1}^\infty A^n$. (This includes showing that the right hand side converges!)
Hint: Recall that $\|AB\| \leq \|A\|\|B\|$.

Solution: As hinted, induction shows that $\|A^n\| \leq \|A\|^n$. So considering the series $\sum_{n=1}^\infty A^n$ as a series in the normed vector space $\mathbb{R}^{m^2}$, which is complete, part (a) shows that the series converges provided we find M_n such that $\|A^n\| \leq M_n$ and such that $\sum_{n=1}^\infty M_n$ converges; the first line of the proof shows that if we take $M_n = \|A\|^n$, we have $\|A^n\| \leq M_n$; and as $\|A\| < 1$, $\sum_{n=1}^\infty M_n$ converges since it is a geometric series with common ratio $\|A\| < 1$. Correspondingly, $\sum_{n=1}^\infty A^n$ converges.

Now, $(I - A)(I + \sum_{n=1}^N A^n) = I - A^{N+1}$, and $\|A^{N+1}\| \leq \|A\|^{N+1}$, which tends to 0 as $N \to \infty$ since $\|A\| < 1$, so $(I - A)(I + \sum_{n=1}^N A^n) \to I$. Since matrix multiplication is a continuous, thus sequentially continuous, map, i.e. $BC_n \to BC$ if $C_n \to C$, we have $(I-A)(I+\sum_{n=1}^N A^n) \to (I-A)(I+\sum_{n=1}^\infty A^n)$. Combining these two, $(I-A)(I+\sum_{n=1}^\infty A^n) = I$, i.e. $I + \sum_{n=1}^\infty A^n$ is a right inverse for $I - A$. A similar computation shows that it is also a left inverse, completing the proof.

6(a) (3 points): Let I be the identity operator on $\mathbb{R}^n$. Show that if $A : \mathbb{R}^n \to \mathbb{R}^n$ is linear then the statements '$AB = I$ for some $B : \mathbb{R}^n \to \mathbb{R}^n$ linear' and '$CA = I$ for some $C : \mathbb{R}^n \to \mathbb{R}^n$ linear' are equivalent, and necessarily $B = C$ in either case.

Solution: Notice that $AB = I$ for some $B : \mathbb{R}^n \to \mathbb{R}^n$ linear means that A is surjective: given any $x \in \mathbb{R}^n$, $x = A(Bx)$. On the other hand, $CA = I$ for some $C : \mathbb{R}^n \to \mathbb{R}^n$ linear means that A is injective, for if $Ax = 0$ then $x = Ix = CAx = 0$. Now, for a linear operator A between finite dimensional vector spaces of the same dimension, injectivity and surjectivity are equivalent by the rank-nullity theorem, i.e. either implies bijectivity, and thus the invertibility of A, with the inverse being a linear map. Thus, under either assumption, A is invertible, with inverse A^{-1}. Since an invertible operator has a unique left inverse, namely A^{-1}, and similarly for a right inverse, we conclude that $A^{-1} = B$, resp. $A^{-1} = C$ in the two cases, and thus in the first case B is the unique left inverse as well so $CA = I$ holds only for $C = B$, with the argument in the second case being similar.

(b) (5 points extra credit only): Consider the set $\mathrm{O}(n)$ of $n \times n$ matrices A with real entries such that $A^{\mathrm{T}}A = I$. (i) Show that $\mathrm{O}(n)$ is a group under matrix multiplication. (ii) Show that there is an open set V in $\mathbb{R}^{n^2}$ containing I such that $\mathrm{O}(n) \cap V$ is a C^1 submanifold of $\mathbb{R}^{n^2}$. What is its dimension?

Note for (ii): In fact, $\mathrm{O}(n)$ is a C^1, and indeed C^∞, submanifold of $\mathbb{R}^{n^2}$, but you do not need to show it.

Hint for (i): Recall that matrix multiplication is associative, and I is a unit for this, so you need to show that $\mathrm{O}(n)$ is closed under multiplication, inverses and contains I. Hint for (ii): Consider the map $A \mapsto \{(A^{\mathrm{T}}A)_{ij} : i \le j\}$ into the above diagonal (including diagonal) entries of the symmetric matrix $A^{\mathrm{T}}A$.

Solution: (i) Note first that if $A^{\mathrm{T}}A = I$, then by part (a) A is invertible, so $\mathrm{O}(n) \subset \mathrm{GL}(n)$, and the latter is a group under matrix multiplication. So it suffices to show that $\mathrm{O}(n)$ is closed under the group operations and contains I; the latter is immediate, and if $A, B \in \mathrm{O}(n)$ then $(AB)^{\mathrm{T}}(AB) = B^{\mathrm{T}}A^{\mathrm{T}}AB = B^{\mathrm{T}}B = I$ so $AB \in \mathrm{O}(n)$, and $AA^{\mathrm{T}} = I$ (as $A^{\mathrm{T}} = A^{-1}$ as shown in (a)), so $(A^{\mathrm{T}})^{\mathrm{T}}A^{\mathrm{T}} = I$, so $A^{\mathrm{T}} \in \mathrm{O}(n)$, completing the proof. (ii) Consider the map $f : R^{n^2} \to \mathbb{R}^{n(n+1)/2}$, with the latter identified with the above diagonal elements of a symmetric matrix, given by $f(A)_{ij} = (A^{\mathrm{T}}A - I)_{ij}$, $i \le j$. Notice that as $A^{\mathrm{T}}A - I$ is symmetric, $f(A) = 0$ if and only if $A^{\mathrm{T}}A - I = 0$, i.e. $\mathrm{O}(n) = f^{-1}(0)$. Let us compute $(Df)(I) : \mathbb{R}^{n^2} \to \mathbb{R}^{n(n+1)/2}$. If we show that this is surjective, i.e. has rank $n(n+1)/2$, the implicit function theorem guarantees that there is an open set V in $\mathbb{R}^{n^2}$ containing I such that $\mathrm{O}(n) \cap V$ is a C^1 submanifold of $\mathbb{R}^{n^2}$ of dimension $n^2 - n(n+1)/2 = n(n-1)/2$. Now, since we are interested in A near I, write $A = I + B$, $B = A - I$. Then $A^{\mathrm{T}}A - I = (I + B^{\mathrm{T}})(I + B) - I = B^{\mathrm{T}} + B + B^{\mathrm{T}}B$. Notice that $\|B^{\mathrm{T}}B\| \le \|B\|^2$, so for the purposes of computing $(Df)(I)$, $B^{\mathrm{T}}B$ can be dropped, and we conclude that $((Df)(I)B)_{ij} = (B^T + B)_{ij}$, $i \le j$. But this is certainly a surjective map: taking B to have ij entry 1 and every other entry 0 gives $((Df)(I)B)_{ij} = 1$ and every other entry $k\ell$ with $k \le \ell$ vanish, so the range of $(Df)(I)$ includes the standard basis of $\mathbb{R}^{n(n+1)/2}$, completing the proof of surjectivity, and thus via the implicit function theorem, the proof.

.

.