Woodbury Matrix Identity	$(A+UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}$
Sherman Morrison Formula	Given *invertible* $A \in R^{n\times n}$ and $u, v \in R^n$ vectors. Then $A + uv^T$ is invertible iff $1 + v^T A^{-1}u \neq 0$. Then $(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^T A^{-1}}{1+v^T A^{-1}u}$
Cauchy Schwarz Inequality	$\lvert\langle U, V\rangle\rvert \leq \langle U, U\rangle \cdot \langle V, V\rangle$
Markov Inequality	X is a non-negative rv and $a > 0$ then $P(X \geq a) \leq \frac{E(X)}{a}$
Minkowski inequality	$\\|f + g\\|_p \leq \\|f\\|_p + \\|g\\|_p$
Holder Inequality	Let $p, q \in [1, \infty], 1/p + 1/q = 1$. Then for all real/complex functions $f, g, \\|fg\\|_1 \leq \\|f\\|_p \\|g\\|_q$
Jensen Inequality	Given X is a rv and ϕ is convex. Then $\phi(E(X)) \leq E(\phi(X))$
Johnson Lindenstrauss Property	A distribution on matrices $S \in R^{k\times n}$ has (ϵ, δ, l)-JL moment property if $\forall x \in R^n, \lvert x\rvert_2 = 1$, $E_S(\lvert\lvert Sx\rvert_2^2 - 1\rvert^l) \leq \epsilon^l \delta$
Hoeffding Inequality	Let $X_1, \cdots, X_n$ be independent rv in $[0, 1]$. Let $\bar{X} = \frac{1}{n}\sum_i X_i$. Then $P(\bar{X} - E(\bar{X}) \geq t) \leq \exp(-2nt^2)$
Chebyshev Inequality	$P(\lvert X - E(X)\rvert \geq a) \leq \frac{Var(X)}{a^2}$
Golden Thompson	For Hermitian matrices A,B $tr \exp(A + B) \leq tr(\exp(A) \exp(B))$
Oppenheimer Inequality	For PSD matrices A, B, $\det(A \circ B) \geq (\prod a_{ii}) \det(B)$
Ky Fan Matrix Inequality	Let A, B be $n \times n$ Hermitian matrices. Then $\lambda(A) + \lambda(B) \succ \lambda(A + B)$, where $\succ$ denotes the majorization relation.
McDiarmid Inequality	Multidimensional Hoeffding. Let $X_1, ...X_m$ be independent rv taking values in χ. Further, let $f : \chi^m \mapsto \mathbb{R}$ be a function of $X_1, ..., X_m$ that satisfies coordinate-wise bounded differences, $\forall i \forall x_1, ...x_m, x_i^* \in \chi$, $\lvert f(x_1, ..., x_i, ...x_m) - f(x_1, ..., x_i^*, ...x_m)\rvert \leq c_i$. Then $\mathbb{P}(\lvert f - \mathbb{E}(f)\rvert \geq \epsilon) \leq \exp\left(\frac{-2\epsilon^2}{\sum_{i=1}^m c_i^2}\right)$.
Cortes Sampling Concentration Inequality	Let $X_i, i \in [m]$ be a set of drawn *without replacement* from an underlying finite set of $m + u$ elements. Let $f : \chi^m \mapsto \mathbb{R}$ be a function of $X_1, ..., X_m$ be symmetric (up to permutation of parameters) and obey coordinate-wise bounded differences, then $\mathbb{P}(\lvert f - \mathbb{E}(f)\rvert \geq \epsilon) \leq \exp\left(\frac{-2\epsilon^2}{\alpha(m,u)c^2}\right)$, where $\alpha(m, u) = \frac{mu}{m+u-0.5} \times \frac{1}{1-1/(2\max(m,u))}$.
Simple Gaussian Concentration Inequality	Let $X_1, ...X_n$ be iid $(0, 1)$ Gaussian rv. Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be 1-Lipschitz. Then for any $\lambda > 0$, $\mathbb{P}(\lvert f(X) - \mathbb{E}(f(X))\rvert \geq \lambda) \leq C \exp(-c\lambda^2)$ for constants c, C.
Bounds for central term in binomial coefficient	$\frac{4^n}{2n+1} \leq \binom{2N}{N} \leq 4^n$

Generating function of Catalan Numbers	$G(z) = \sum_{i=0}^{\infty} \frac{1}{n+1}\binom{2n}{n} z^i = \frac{1-\sqrt{1-4z}}{2z} = \frac{2}{1+\sqrt{1-4z}}$
Weyl Matrix Perturbation Inequality	If $M = H + P$ are $n \times n$ matrices with eigenvalues μ_i, ν_i, ρ_i (ordered in descending order) respectively, then $\nu_i + \rho_n \leq \mu_i \leq \nu_i + \rho_1$
Walds Identity	Let $X_1, ...$ be a sequence of (potentially infinite) independent random variables with identical means. Let N be any *stopping time* (or independent of Xs). Let $S_N = \sum_{i=1}^{N} X_i$. Then $\mathbb{E}(S_N) = \mathbb{E}(N)\mathbb{E}(X)$. Note N can be weakly dependent on X.
Variance of Random Sum	If N and X_i are independent and Xs have equal variance, $\text{Var}(\sum_{i=1} X) = \mathbb{E}(N)\text{Var}(X_i) + \mathbb{E}(X)^2\text{Var}(N)$
Eckart-Young-Mirsky Theorem	The k-truncated SVD of a matrix A gives the best rank-k approximation for A in the spectral and frobenius norm. $A_k = U\Sigma_k V^T = argmin_{\text{B rank-k}} \lvert A - B \rvert_2 = \sigma_{k+1}(A)$
Wolfe Conditions	(i) (Armijo rule) $f(\mathbf{x}_k + \alpha_k \mathbf{p}_k) \leq f(\mathbf{x}_k) + c_1 \alpha_k \mathbf{p}_k^{\mathrm{T}} \nabla f(\mathbf{x}_k)$ (ii) (Curvature) $-\mathbf{p}_k^{\mathrm{T}} \nabla f(\mathbf{x}_k + \alpha_k \mathbf{p}_k) \leq -c_2 \mathbf{p}_k^{\mathrm{T}} \nabla f(\mathbf{x}_k)$, where α_k is step size, $\mathbf{p}_k$ is descent direction and $0 < c_1 < c_2 < 1$